



# Agencja Oceny Technologii Medycznych i Taryfikacji



**Fundusze Europejskie**  
Wiedza Edukacja Rozwój



**Rzeczpospolita  
Polska**

**Unia Europejska**  
Europejski Fundusz Społeczny



## **Podejście światowych agencji HTA do istotności klinicznej punktów końcowych**

The approach of global HTA agencies to clinical relevance of endpoints

Tomasz Tatar<sup>1</sup>, Justyna Misiak<sup>1</sup>, Marcin Rędownicz<sup>1</sup>

<sup>1</sup>Agencja Oceny Technologii Medycznych i Taryfikacji,  
Agency for Health Technology Assessment and Tariff System

Warszawa, sierpień 2019

## Spis treści

<b>Spis treści .....</b>	<b>2</b>
<b>1 Abstrakt.....</b>	<b>3</b>
<b>2 Wstęp.....</b>	<b>4</b>
2.1 Cel.....	5
2.2 Metodyka .....	5
<b>3 Wyniki.....</b>	<b>6</b>
3.1 Anglia.....	6
3.2 Australia.....	7
3.3 Belgia.....	8
3.4 Chorwacja .....	9
3.5 Czechy .....	9
3.6 Finlandia .....	10
3.7 Francja .....	10
3.8 Holandia.....	10
3.9 Irlandia .....	12
3.10 Kanada .....	14
3.11 Niemcy.....	14
3.12 Norwegia.....	15
3.13 Nowa Zelandia.....	16
3.14 Szkocja.....	16
3.15 Szwecja.....	17
3.16 Węgry.....	17
3.17 EUnetHTA .....	18
<b>4 Podsumowanie i wnioski.....</b>	<b>23</b>
<b>5 Bibliografia .....</b>	<b>24</b>

## 1 Abstrakt

Celem pracy było zbadanie podejścia wybranych instytucji zajmujących się oceną technologii medycznych do istotności klinicznej punktów końcowych w oparciu o publikowane, powszechnie dostępne dokumenty (wytyczne, zalecenia, stanowiska). Przeszukano strony internetowe wybranych instytucji związanych z oceną technologii medycznych (HTA), będących członkami Europejskiej sieci agencji oceny technologii medycznych (European Network for Health Technology Assessment, EUnetHTA) oraz będących członkami międzynarodowego towarzystwa ISPOR (The International Society for Pharmacoeconomics and Outcomes Research). W wyniku wyszukiwania odnaleziono wytyczne agencji i instytucji związanych z HTA z 16 państw. W większości wytycznych omawiano punkty końcowe w kontekście oceny analizy ekonomicznej. Wskazuje się, że wybór konkretnego punktu końcowego zależy od takich czynników jak: populacja pacjentów spełniających kryteria kwalifikacji, stan zdrowia i cel leczenia. Punkty końcowe powinny odzwierciedlać przydatność interwencji w zakresie opieki zdrowotnej. Spośród punktów końcowych definiowanych jako istotne klinicznie najczęściej wymieniano te: związane z przeżyciem, przebiegiem/nasileniem choroby, jakością życia związaną ze zdrowiem oraz występowaniem zdarzeń/działania niepożądanych. Zastosowanie zastępczych punktów końcowych (surogatów) jest zasadne, gdy: nie jest możliwe zastosowanie klinicznie istotnych punktów końcowych oraz wykazano wpływ tego punktu na klinicznie istotne punkty końcowe. Wykorzystanie złożonych punktów końcowych jest zasadne, gdy uwzględnione składowe złożonego punktu końcowego są istotnymi punktami końcowymi powiązanymi klinicznie. Uzasadnieniem łączenia różnych miar wyników jest zwiększenie mocy statystycznej dla punktów końcowych, których częstotliwość występowania jest niska. Odnalezione wytyczne zagranicznych agencji i instytucji związanych z oceną technologii medycznych w zakresie definicji istotnych punktów końcowych, zastosowania zastępczych punktów końcowych oraz wykorzystania złożonych punktów końcowych są spójne z wytycznymi polskiej agencji HTA (AOTMiT) oraz EUnetHTA.

Słowa kluczowe: wytyczne, HTA, skuteczność/efektywność kliniczna, istotność kliniczna, kliniczne punkty końcowe, surogaty, złożone punkty końcowe.

The goal of the presented paper is to analyse the approach of selected health technology assessment institutions towards the clinical significance of outcomes, basing on published and broadly accessible documents (i.e. guidelines, recommendations, positions). The websites of selected institutions involved in health technology assessment (HTA) united in the European Network for Health Technology Assessment (EUnetHTA) and members of The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) were searched. As a result, guidelines from HTA agencies and institutions from 16 countries were identified. In most of the guidelines, outcomes were discussed in the context of economic analysis. This findings that a choice of particular outcome depends on factors such as: patients population fulfilling inclusion criteria, health condition, and the purpose of a treatment. Outcomes should reflect the usefulness of interventions in a healthcare system. From the endpoints defined as clinically significant the most often mentioned were related to: survival, the morbidity, health related life quality, and the occurrence of side effects. Use of surrogate outcomes is justified when the following criteria are met: it is impossible to use a clinically significant outcomes and the influence of a surrogate outcome on a clinically significant outcomes has been demonstrated. Employing surrogate outcomes is only reasonable when the included elements are clinically related to significant outcomes. The rationale for combining different measures is to increase the statistical power of seldom occurring endpoints. The identified guidelines of agencies and HTA institutions regarding the definition of significant outcomes, the use of surrogate outcomes and the use of complex outcomes are consistent with the guidelines of the Polish HTA Agency (AOTMiT) and the EUnetHTA.

Keywords: guidelines, HTA, efficacy, clinical significance, clinical outcomes, surrogates, complex outcomes.

## 2 Wstęp

Do zadań Agencji Technologii Medycznych i Taryfikacji (AOTMiT), zgodnie z art. 31n ustawy z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych, należy m.in.:

- wydawanie rekomendacji w sprawie świadczeń gwarantowanych;
- opracowywanie raportów w sprawie oceny świadczeń opieki zdrowotnej;
- opracowywanie analiz weryfikacyjnych, o których mowa w art. 35 ustawy o refundacji, które Minister właściwy do spraw zdrowia bierze pod uwagę, wydając decyzję administracyjną o objęciu refundacją i ustaleniu urzędowej ceny zbytu leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych (1,2).

W celu wskazania zasad i akceptowalnych metod przeprowadzania oceny technologii medycznych (ang. Health technology assessment, HTA), zapewniających wysoką jakość analiz i wiarygodność ich wyników w 2016 roku Agencja wydała zaktualizowane wytyczne oceny technologii medycznych.

Według wytycznych proces HTA należy rozpocząć od określenia kierunku i zakresu analiz oraz metody, w tym zasady wyboru danych i informacji zawartych w raporcie HTA, które powinny być przedstawione w analizie problemu decyzyjnego (APD). APD stanowi przegląd podstawowych informacji niezbędnych do poprawnego wykonania raportu HTA, pozwala prawidłowo zbudować kryteria włączania badań do analizy według schematu PICOS (ang. population, intervention, comparison, outcome, study):

- populacja, w której dana interwencja będzie stosowana (P);
- proponowana interwencja (I);
- proponowane komparatory (C);
- efekty zdrowotne, czyli punkty końcowe, względem których oceniana będzie efektywność kliniczna (O);
- rodzaj włączanych badań (S) (3).

Ważnym elementem przedstawionego powyżej schematu PICOS są punkty końcowe.

Polski Instytut Evidence Based Medicine definiuje punkt końcowy jako określoną zmianę w stanie zdrowia (np. zawał serca, zgon, ustąpienie bólu itp.) mogącą zajść u osoby uczestniczącej w badaniu, której wystąpienie albo niewystąpienie może się wiązać z ekspozycją na określony czynnik lub interwencję. Na podstawie różnicy w częstości występowania punktów końcowych w badanych grupach porównuje się skuteczność i bezpieczeństwo różnych interwencji (np. leków lub zabiegów) (4).

Ocena korzyści zdrowotnych wnoszonych przez ocenianą technologię medyczną, zgodnie z wytycznymi AOTMiT, powinna być dokonywana poprzez analizę istotnych klinicznie punktów końcowych, odgrywających kluczową rolę w danej jednostce chorobowej. Wskazuje się cztery główne kategorie, grupujące istotnie kliniczne punkty końcowe: punkty końcowe odnoszące się do śmiertelności, przebiegu/nasilenia choroby oraz zależnej od zdrowia jakości życia (ang. health-related quality of life, HRQoL), a także punkty końcowe dotyczące zdarzeń i działań niepożądanych.

Wskazuje się, iż punkty końcowe raportowane w analizach powinny być zdefiniowane i uzasadnione w opisie problemu decyzyjnego; dotyczyć ocenianej jednostki chorobowej i jej przebiegu; odzwierciedlać medycznie istotne aspekty problemu zdrowotnego i jednocześnie umożliwiać wykrycie potencjalnych różnic między porównywanymi technologiami oraz mieć zasadnicze znaczenie dla podejmowania racjonalnych decyzji klinicznych (punkty krytyczne dla danego problemu zdrowotnego).

Wyniki leczenia należy poddać analizie w najdłuższym dostępnym okresie obserwacji. Ocena wyników leczenia w krótkim czasie obserwacji jest wystarczająca w problemach zdrowotnych o przebiegu ostrym, bez konsekwencji długoterminowych. W chorobach przewlekłych wyższą wartość mają wyniki uzyskane w dłuższym okresie obserwacji, jakkolwiek w niektórych sytuacjach ocena skuteczności leczenia – ze względu na długie oczekiwane przeżycie – będzie musiała być dokonana w oparciu o wyniki uzyskane w krótszym okresie obserwacji.

Wytyczne zalecają, jeśli ocena efektywności klinicznej jest przeprowadzona w oparciu o wyniki w zakresie zastępczych (surogatowych) punktów końcowych, aby w analizie klinicznej wiarygodnie wykazać ich związek z klinicznie istotnymi punktami końcowymi. Walidacja surogatowych punktów końcowych powinna być dokonana w odniesieniu do rozpatrywanego problemu zdrowotnego.

Stosowanie złożonych punktów końcowych zalecane jest wyłącznie w przypadku, gdy zostały one predefiniowane w protokole badania klinicznego. Nie zaleca się analizowania złożonych punktów końcowych zdefiniowanych

w ramach analizy post-hoc; jeśli analizy w podgrupach chorych były predefiniowane na etapie planowania badania, ich wyniki mają większe znaczenie niż typowych analiz post-hoc. W przypadku raportowania złożonych punktów końcowych niezbędne jest podanie nie tylko wyników dla złożonego punktu końcowego, ale też oddzielnie dla każdego komponentu, nawet gdy nie osiągnęły istotności statystycznej.

Podkreśla się także, że gdy wyniki oceny klinicznej uzyskane są przy użyciu skal lub kwestionariuszy, należy przedstawić informacje o ich walidacji oraz istotności klinicznej wyników. W przypadku konwersji zmiennych ciągłych lub porządkowych na zmienne dichotomiczne (np. zdrowy – chory) konieczne jest uzasadnienie punktu odcięcia (3).

Biorąc pod uwagę kluczowy wpływ uwzględnionych w HTA punktów końcowych postanowiono zbadać podejście światowych agencji HTA do istotności klinicznej punktów końcowych.

### **2.1 Cel**

Celem pracy było zbadanie podejścia wybranych instytucji zajmujących się HTA do istotności klinicznej punktów końcowych w oparciu o publikowane, powszechnie dostępne dokumenty (wytyczne, zalecenia, stanowiska).

### **2.2 Metodyka**

W celu odnalezienia informacji dotyczących podejścia światowych agencji HTA do istotności klinicznej punktów końcowych przeszukano strony internetowe wybranych instytucji związanych z HTA, będących członkami Europejskiej sieci agencji oceny technologii medycznych (European Network for Health Technology Assessment, EUnetHTA) oraz będących członkami międzynarodowego towarzystwa ISPOR (The International Society for Pharmacoeconomics and Outcomes Research).

Wyszukiwanie przeprowadzone zostało przez dwóch niezależnie pracujących analityków (J.M., M.R.) w czerwcu 2019 roku. Na angielskich wersjach językowych stron instytucji przeszukiwano zakładki oraz wykorzystano wyszukiwarki na ww. stronach przy zastosowaniu słów kluczowych: „guideline”, „guidelines”, „HTA”, „clinical efficacy”, „clinical effectiveness”, „clinical outcome”, „assessment”, „endpoint”, „surrogate”. Do pracy włączano dokumenty w języku angielskim opublikowane w przeciągu ostatnich 10 lat.

### 3 Wyniki

W wyniku wyszukiwania odnaleziono łącznie 27 dokumentów z 16 krajów. Dodatkowo przedstawiono treść wytycznych EUnetHTA, w których zaprezentowano informacje na temat istotności klinicznej punktów końcowych. Poniżej przedstawiono najważniejsze informacje z odnalezionych wytycznych.

#### 3.1 Anglia

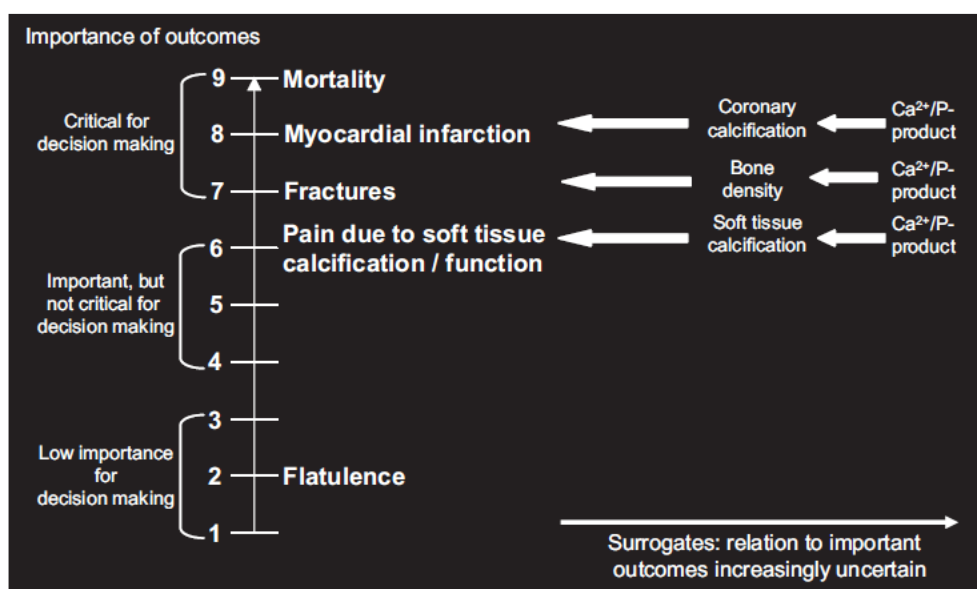
Wytyczne brytyjskiej agencji HTA (National Institute for Health and Care Excellence, NICE) z 2014 roku (aktualizacja w 2018 roku) obejmują wytyczne praktyki klinicznej, praktyki medycznej, zdrowia publicznego i opieki społecznej.

Formułowanie pytania klinicznego odnośnie oceny skuteczności interwencji następuje przy użyciu schematu PICO (ang. population, intervention, comparison, outcome).

Według NICE przy formułowaniu PICO w odniesieniu do punktów końcowych należy uwzględnić m.in. następujące pytania: Jakie punkty końcowe należy rozważyć, aby ocenić działanie interwencji (w tym wyniki dotyczące zarówno korzyści, jak i negatywne efekty)? Co jest szczególnie istotne dla osób korzystających z usług medycznych? Należy stosować podstawowe wyniki zdrowotne, na podstawie oceny jakości i przydatności. Jednym ze źródeł jest baza danych COMET. Standardy COS-STAD (ang. The Core Outcome Set Standards for Development) i COS-STAR (ang. Core Outcome Set Standards for Reporting) pozwalające poprawić jakość analiz HTA powinny być stosowane do oceny przydatności zidentyfikowanych, podstawowych wyników zdrowotnych.

W przypadku każdego pytania klinicznego należy rozważyć czynniki, które mogą mieć wpływ na wyniki i skuteczność interwencji, w tym szersze czynniki społeczne, które mogą wpływać na zdrowie i nierówności zdrowotne. Wyniki dotyczące zarówno korzyści, jak i szkód oraz pozostałe powinny być wstępnie określone w protokole przeglądu. Należy zdefiniować maksymalnie od siedmiu do dziesięciu punktów końcowych. Wytyczne dotyczące priorytetów wyboru punktów końcowych zostały opracowane przez grupę roboczą GRADE (ang. The Grading of Recommendations Assessment, Development and Evaluation) (5).

GRADE wymaga ustalenia jasnej specyfikacji hierarchii pytania klinicznego: populacji, interwencji i komparatora. Wymaga się również określenia wszystkich istotnych wyników zdrowotnych. Grupa GRADE zaproponowała przykładowy schemat hierarchii punktów końcowych dla oceny wpływu leków obniżających poziom fosforanów u pacjentów z niewydolnością nerek i hiperfosfatemią – rysunek poniżej (6).



Rysunek 1. Hierarchia wyników według ich znaczenia dla oceny wpływu leków obniżających poziom fosforanów u pacjentów z niewydolnością nerek i hiperfosfatemią (6)

W rekomendacji z 2015 roku, obejmującej zasady bezpiecznego i skutecznego stosowania leków w opiece zdrowotnej i społecznej dla osób przyjmujących 1 lub więcej leków, przy formułowaniu pytania klinicznego, NICE zaleca rozważyć następujące punkty końcowe:

- optymalnie przypisane do choroby,
- w ramach oceny bezpieczeństwa: zdarzenia związane z leczeniem,
- wyniki zgłaszane przez pacjentów (np. zadowolenie pacjentów i przestrzeganie zaleceń lekarskich),
- jakość życia,
- wyniki istotne klinicznie,
- wykorzystanie zasobów opieki zdrowotnej i społecznej.

W celu zapewnienia rzetelności i wiarygodności, wyniki zdrowotne powinny być mierzone przy użyciu zwalidowanych narzędzi. Tam, gdzie nie jest to możliwe, miara wyniku powinna być szczegółowo opisana w badaniu. Jakość życia należy oceniać za pomocą kwestionariusza EQ-5D (ang. European Quality of Life 5-Dimensions questionnaire), w celu możliwości przeprowadzenia analizy kosztów użyteczności (7).

W wytycznych skierowanych do wnioskodawców, w ramach analizy efektywności klinicznej, NICE zaleca przedstawienie wyników dla wszystkich punktów końcowych, istotnych dla oceny ekonomicznej lub rozważanego problemu (wyniki analizy klinicznej muszą obejmować pierwszorzędowe punkty końcowe) (8).

W kolejnej publikacji dotyczącej procesu i metod zbierania dowodów naukowych dla nowych leków, NICE definiuje punkty końcowe ukierunkowane na pacjenta (takie, które mają bezpośrednie znaczenie kliniczne): śmiertelność, współczynnik zdarzeń sercowo-naczyniowych lub jakość życia. W odróżnieniu do nich surogaty ukierunkowane są na wyniki zdrowotne związane z chorobą, np. zmiana ciśnienia krwi lub badania biochemiczne (9).

W wytycznych dotyczących metodologii HTA, NICE zaleca dobór punktów końcowych określających główne miary wyników zdrowotnych, które będą istotne dla oceny skuteczności klinicznej. Oznacza to, że mierzą one korzyści zdrowotne i niekorzystne skutki zdrowotne, które są istotne dla pacjentów i/lub ich opiekunów. Istotne klinicznie punkty końcowe zwykle określają wpływ na przeżycie lub jakość życia związaną ze zdrowiem, co przekłada się na lata życia skorygowane o jakość (ang. quality-adjusted life year, QALY) mające wpływ na ocenę opłacalności.

Klinicznie istotne punkty końcowe, odzwierciedlające sposób, w jaki czuje się, funkcjonuje lub jak długo przeżywa pacjent, są bardziej przydatne, niż zastępcze punkty końcowe (takie jak: testy laboratoryjne i wyniki obrazowania). Gdy zastosowanie istotnych klinicznie punktów końcowych nie jest możliwe, a dane pochodzące z surogatów są wykorzystywane do wnioskowania o wpływie leczenia na śmiertelność i jakość życia związaną ze zdrowiem, należy dostarczyć dowody na poparcie niniejszej tezy i związku/wpływu surogatów na istotne klinicznie punkty końcowe.

Użyteczność wykorzystania zastępczego punktu końcowego dla oszacowania QALY będzie największa, gdy istnieją mocne dowody wskazujące na wpływ tego punktu na jakość życia związaną ze zdrowiem i/lub przeżycie. We wszystkich przypadkach należy ilościowo zbadać i określić niepewność związaną z zależnością między punktem końcowym, a jakością życia związaną ze zdrowiem lub przeżyciem (10).

### **3.2 Australia**

Wytyczne Pharmaceutical Benefits Advisory Committee (PBAC) z 2016 roku dotyczą przygotowywania wniosków refundacyjnych dla tego komitetu.

Przy definiowaniu schematu PICO punkty końcowe powinny krótko określać istotną dla pacjenta skuteczność i bezpieczeństwo (powinny być związane z jakością i/lub długością życia).

PBAC zaleca, aby podczas charakterystyki każdego badania opisać pierwszorzędowe punkty końcowe, wskazane w protokole badania oraz drugorzędowe punkty końcowe istotne dla pacjenta.

Surogaty (które nie są pierwszorzędowymi punktami końcowymi) przedstawia się tylko, kiedy jest to istotne dla konkluzji terapeutycznych lub oceny ekonomicznej. W analizie ekonomicznej należy uzasadnić i określić ilościowo deklarowany związek między zmianą w efekcie leczenia w pomiarze surogatu, a zmianą w efekcie leczenia w docelowym, klinicznym punkcie końcowym zastosowanym w analizie ekonomicznej.

W obrębie punktów końcowych PBAC wydziela: złożone punkty końcowe, punkty końcowe raportowane przez pacjenta oraz zdarzenia niepożądane.

Złożone punkty końcowe składają się z wielu, połączonych punktów końcowych. Należy przedyskutować i porównać kliniczną ważność każdego z komponentu. Istotne jest, aby wskazać czy definicję złożonego punktu

końcowego predefiniowano. Należy także uzasadnić uwzględnienie komponentów w złożonym punkcie końcowym.

Punkty końcowe raportowane przez pacjenta obejmują generyczne (globalne) lub specyficzne dla choroby pomiary jakości życia, objawów lub funkcji (np. układu oddechowego, depresji, artretyzmu). Pomiary punktów końcowych raportowanych przez pacjenta mogą także obejmować wieloatrybutowe narzędzia użyteczności (ang. multiattribute utility instruments, MAUIs), w których metoda skoringowa dla instrumentu jest zakotwiczona na skali życia skorygowanej o jakość od 0 (zgon) do 1 (pełne zdrowie). Kilkoma powszechnie stosowanymi MAUIs, dla których nie jest wymagane przedstawienie szczegółowej dyskusji ważności lub wiarygodności są: HUI2 lub HUI3 (ang Health Utilities Index mark 2, 3), EQ5D-3L lub -5L ('EuroQol'), SF-6D (podzestaw Short Form 36 – SF-36), Assessment of Quality of Life (AQoL) instruments oraz Child Health Utility 9D (CHU9D) index dla dzieci i młodzieży.

Powinno się raportować co najmniej: jakiegokolwiek zdarzenia niepożądane (ZN), jakiegokolwiek ZN prowadzące do przerwania leczenia w randomizowanym badaniu, jakiegokolwiek ciężkie ZN, jakiegokolwiek ZN prowadzące do zgonu, każde ZN, w którym częstość lub ciężkość różniły się znacznie między grupami w poszczególnym badaniu.

Przy wyborze istotnego punktu końcowego należy się upewnić, że jest on znaczący dla pacjenta i zawiera kluczowy cel interwencji w procesie chorobowym. Celem leczenia może być poprawa jakości życia lub zapobieganie bądź spowolnienie choroby długoterminowej.

PBAC preferuje wnioski refundacyjne, które nie opierają się na surogatach. Jeżeli jest to możliwe powinno się przedstawiać dowody naukowe z bezpośrednich badań randomizowanych oceniających efekt leczenia danego leku dla klinicznie istotnych punktów końcowych. Gdy nie ma takich dowodów, ustala się prawdopodobny, porównawczy efekt leczenia dla klinicznie istotnego punktu końcowego poprzez przekształcenie efektu leczenia surogatu. Miarą surogatu jest biomarker, który ma zastąpić jeden lub więcej klinicznie istotnych punktów końcowych.

Gdy istnieje duża, niepewność oszacowania punktu końcowego, który występuje przed zmianą (ang. switching) leczenia, np. przeżycie wolne od progresji choroby (ang. progression free survival, PFS), należy przedyskutować czy jest on ważnym surogatem dla klinicznie istotnego punktu końcowego, np. przeżycia całkowitego (ang. overall survival, OS). Kiedy PFS jest uzasadnionym surogatem dla OS, należy porównać oszacowanie OS przez przekształcenie PFS z OS określoną metodą statystyczną, zastosowaną do dopasowania przełączania leczenia.

Jednym z komponentów analizy ekonomicznej są punkty końcowe, np. uniknięcie zdarzeń, zyskane lata życia, lata życia skorygowane o jakość. PBAC preferuje analizę użyteczności kosztów (ang. cost-utility analysis, CUA) niż analizę efektywności kosztów (ang. cost-effectiveness analysis, CEA), gdy:

- można założyć, że uzyskano dodatkowe lata życia w ocenie ekonomicznej (żeby ocenić wpływ jakości dopasowanej do zyskanych lat życia),
- uzyskano poprawę w jakości, ale nie w długości życia lub
- bezpośrednio, randomizowane badania raportowały wyniki z zastosowaniem wieloatrybutowych instrumentów oceny użyteczności.

W QALY powinno się także wyrazić kombinację punktów końcowych (pośrednich lub ostatecznych, bądź jednych i drugich), które są istotne dla pacjenta. W analizie kosztów-efektywności należy się upewnić, że inkrementalne punkty końcowe (np. lata życia, inne zdarzenia zdrowotne) są istotne dla pacjenta. Należy także uzasadnić niemożliwość wyrażenia punktów końcowych w QALY (11).

### 3.3 Belgia

Wytyczne Health Care Knowledge Centre (KCE) z 2012 roku dotyczą oceny ekonomicznej i analizy wpływu na budżet.

Dobry przegląd rozpoczyna się identyfikacją pytania klinicznego, które uwzględnia, m.in. punkty końcowe. Warte uwagi są punkty końcowe specyficzne dla choroby, zdarzenia niepożądane, przeżycie i jakość życia.

Wyniki oceny ekonomicznej powinny być wyrażone w ostatecznych punktach końcowych, a nie pośrednich. Ważne jest, aby uwzględnić konsekwencje punktów końcowych, w tym związane z pozytywnym i negatywnym wpływem leczenia (np. działania niepożądane).

CEA powinna być wykonana, jeżeli poprawa spodziewanej długości życia jest głównym celem terapii, a także najważniejszym punktem końcowym dla pacjenta. CUA powinna być wykonana, jeżeli leczenie ma wpływ



na jakość życia, która jest istotna dla pacjenta lub istnieje wiele parametrów istotnych dla pacjenta wyrażonych w różnych jednostkach, które nie mogą być zmienione w jedną, powszechną jednostkę we właściwy sposób. Ostatecznymi, zdrowotnymi punktami końcowymi w CEA są zyskane lata życia dla interwencji z wpływem na śmiertelność, a w analizie kosztów-żyteczności – QALY z jakością życia mierzoną na podstawie empirycznych danych, uzyskanych generycznym instrumentem, takim jak: EQ-5D, dla którego istnieją wartości preferencji dla populacji generalnej.

CUA zawsze powinna być uzupełniona o CEA, jeżeli leczenie ma wpływ na jakość życia, która jest istotna dla pacjentów lub leczenie jest związane z wieloma klinicznymi punktami końcowymi, które są wyrażone w różnych jednostkach (np. działania niepożądane vs przeżycie całkowite).

Analizę minimalizacji kosztów (CMA) stosuje się, gdy efekty dwóch terapii są identyczne (12).

### 3.4 Chorwacja

Wytyczne Agencji Jakości i Akredytacji w Ochronie Zdrowia (Agency for Quality and Accreditation in Health Care) z 2011 roku dotyczą procedur i dokumentowania przebiegu oceny technologii medycznych.

Z punktu widzenia Agencji, najbardziej odpowiednie dla procesu oceny i tworzenia rekomendacji dla decydentów jest uwzględnienie wszystkich bezpośrednich wyników dotyczących zdrowia pacjenta. W razie potrzeby w ocenie należy również uwzględnić wpływ na zdrowie innych osób (głównie opiekunów). Efekty te należy interpretować odrębnie od głównych wyników analizy (tj. opisowo), chyba że istnieją wysokiej jakości, wiarygodne dane dotyczące skutków zdrowotnych u nieformalnych opiekunów (z ang. informal caregivers). Zamiar włączenia takich danych powinien być standardowo uzgadniany z Agencją przed zakończeniem przekazywania dokumentacji.

Celem analizy skuteczności klinicznej jest uzyskanie obiektywnego oszacowania średniej skuteczności klinicznej porównywanych technologii. Analiza skuteczności klinicznej powinna opierać się na danych ze wszystkich odpowiednich badań o jak najlepszej jakości i powinna uwzględniać zakres typowych pacjentów, standardowe okoliczności kliniczne, klinicznie istotne wyniki, porównanie z odpowiednimi komparatorami oraz względne i bezwzględne miary z właściwymi pomiarami niepewności.

W wytycznych wskazano także, iż przy zastosowaniu w analizie ekonomicznej techniki CEA, wszystkie bezpośrednie skutki zdrowotne należy wyrazić w kategoriach naturalnych jednostek, takich jak: liczba unikniętych zgonów, zmniejszona częstość powikłań, redukcja skutków ubocznych itp. W analizie żyteczności kosztów wynik powinien być wyrażony QALY. Pomiar zmian wyników jakości życia powinien być przeprowadzany bezpośrednio na pacjentach (lub ich opiekunach), natomiast użyteczność przypisana tej zmianie powinna pochodzić z zestawu wartości użyteczności uzyskanego poprzez pomiar preferencji różnych stanów zdrowia w populacji Chorwacji (13).

### 3.5 Czechy

Dokument Państwowego Instytutu ds. Kontroli Leków z 2017 roku (State Institute for Drug Control) obejmuje procedurę krytycznej oceny analizy kosztów efektywności.

Do celów procedury pojęcie miary wyniku definiuje się, jako parametr korzyści z leczenia wygenerowany przez zastosowanie interwencji badanej i porównawczej, jest on wspólny dla obu tych grup i jest istotny dla wyniku oceny farmakoekonomicznej.

W określonych przypadkach, gdy korzyści z interwencji objętej oceną i interwencji porównawczej są porównywalne w oparciu o dostępne dowody, autorzy analizy powinni wybrać CMA. Porównywalny poziom korzyści musi być wyraźnie poparty odpowiednimi dowodami, poprzez bezpośrednie badanie porównawcze odpowiedniej mocy zaprojektowane w celu oceny, czy badana interwencja nie jest gorsza (ang. non-inferiority study) lub równoważna (ang. equivalence study) względem głównych parametrów skuteczności klinicznej i bezpieczeństwa lub metaanalizy dostępnych badań klinicznych albo innego rodzaju odpowiednie dowody.

Gdy korzyści z ocenianej interwencji i komparatora są różne, a także, gdy rozważana jednostka chorobowa i jej powikłania oraz zastosowane leczenie wpływa na oczekiwaną długość życia lub jego jakość, należy zastosować analizę CUA, gdzie korzyści wyrażone są w postaci QALY.

Tylko w odpowiednio uzasadnionych przypadkach, w których nie można przeprowadzić CUA, dopuszczalne jest zastosowanie analizy CEA, z korzyścią wyrażoną w postaci zyskanych lat życia (ang. life years gained, LYG) lub tam, gdzie nie można zastosować LYG, możliwe jest zastosowanie istotnego i wymiernego kryterium oceny dla danej choroby.

Hierarchia siły i istotności dowodów oraz wymagania dla określonych rodzajów badań są regulowane międzynarodowymi standardami (w dokumencie wskazano wytyczne EUnetHTA oraz NICE) (14).

### 3.6 Finlandia

Dokument z 2017 roku stanowi wytyczne fińskiej Rady ds. Cen Leków (Lääkkeiden hintalautakunta, Läkemedelsprismnämnden, HILA) dla wnioskodawców dotyczy przygotowania analizy ekonomicznej.

Według Rady ds. Cen Leków skuteczność powinna być mierzona głównie w QALY, które obliczono za pomocą zwalidowanej ogólnej miary jakości życia. Skuteczność może być również mierzona na przykład poprzez istotne punkty końcowe (ang. final endpoints), zastępcze punkty końcowe lub specyficzną dla choroby jakość życia. Wybór punktów końcowych powinien być uzasadniony.

Wnioskodawca powinien wziąć pod uwagę, że w szczególności w modelach z długim horyzontem czasowym, niepewność modelowania może być związana z wykorzystaniem zastępczych punktów końcowych. Niepewność oszacowań należy ocenić za pomocą analizy wrażliwości (15).

### 3.7 Francja

Wytyczne francuskiej agencji HTA (Haute Autorité de Santé) z 2012 roku dotyczą wyboru techniki analitycznej analizy ekonomicznej.

Wybór metody analizy ekonomicznej (CUA i/lub CEA) zależy od rodzaju oczekiwanych efektów zdrowotnych badanych interwencji.

Jeżeli interwencja ma istotny wpływ na HRQoL (ang. health related quality of life), należy zastosować CUA. Wynikiem leczenia jest długość życia pacjenta ważona wartością HRQoL. CUA zawsze towarzyszy analiza opłacalności, która wykorzystuje długość życia jako wynik zdrowotny.

Zaleca się stosowanie systemów klasyfikacji stanu zdrowia, dla których potwierdzone wyniki zdrowotne oparte na preferencjach są dostępne we Francji. W momencie pisania wytycznych dostępne były EQ-5D i HUI3.

Jeśli jakość życia związana ze zdrowiem nie jest identyfikowana jako istotny efekt zdrowotny badanych interwencji, CEA jest wymaganą formą oceny ekonomicznej, a wyniki zdrowotne mierzone są długością życia. Niemożność wykorzystania długości życia jako wyniku zdrowotnego w analizie efektywności kosztowej, a także wybór kryterium innego niż lata życia (ang. life years, LY) musi być należycie uzasadnione.

Modele decyzyjne są szczególnie odpowiednie do oceny ekonomicznej zdrowia. Porównanie interwencji zdrowotnych opartych na kryterium efektywności wymaga integracji różnych rodzajów informacji uzyskanych z różnych źródeł (klinicznych, ekonomicznych, epidemiologicznych, socjologicznych, biologicznych, technologicznych itp.). Dane te mogą pochodzić z różnych typów badań modelowania, takich jak model zależności między zastępczymi i istotnymi klinicznie punktami końcowymi (np. poziom HbA1c i powikłania cukrzycy), metaanaliz itp. (16).

### 3.8 Holandia

Wytyczne holenderskiej agencji HTA (Zorginstituut Nederland, ZIN) z 2016 roku dotyczą analizy ekonomicznej.

W badaniach klinicznych, efektywność analizowanej interwencji w porównaniu ze standardową opieką, określana jest za pomocą wcześniej zdefiniowanych miar wyniku (punktów końcowych). Można wyróżnić różne miary wyników. Miary wyników klinicznych odzwierciedlają chorobowość, śmiertelność lub jakość życia. Miary zgłaszane przez pacjentów (patient-reported outcome measures, PROM) odnoszą się do wyników bezpośrednio zgłaszanych przez samych pacjentów, takich jak ból lub jakość życia. Czasami konieczne jest zastosowanie pośrednich miar wyniku, w przypadku gdy w oparciu o badanie niskiej wiarygodności nie jest możliwe wykazanie różnic dla twardych punktów końcowych (np. śmiertelność). Warunkiem wstępnym dla pomiarów pośrednich lub zastępczych punktów końcowych jest udowodniona wrażliwość tych punktów na zmianę istotnego klinicznie punktu końcowego. Inną kwestią jest dokładność analityczna testu, podczas którego mierzony jest pośredni lub zastępczy punkt końcowy. Złożone miary wyników można stosować, gdy rozważa się dwa wyniki, które są wyraźnie powiązane klinicznie. Są one często wykorzystywane do wyników, których częstość występowania jest niska (np. śmiertelność). Innym złożonym punktem jest QALY, który łączy lata życia z jakością życia.

Wybór konkretnego punktu końcowego zależy od takich czynników jak: populacja pacjentów spełniających kryteria kwalifikacji, stan zdrowia i cel leczenia. Zaleca się stosowanie zatwierdzonych/zwalidowanych miar wyniku (wewnętrzna walidacja, odtwarzalność i czułość pomiaru punktów końcowych opiera się na publikacjach naukowych), które są istotnie związane ze stanem zdrowia i celem leczenia. Punkty końcowe powinny

odzwierciedlać przydatność interwencji w zakresie opieki zdrowotnej. Dlatego ważne jest, aby wybrać je przed rozpoczęciem badania, w porozumieniu z grupami zawodowymi, opiekunami i organizacjami pacjentów.

QALY jest standardową miarą wyniku zdrowotnego w analizach ekonomicznych opieki zdrowotnej. Na wyniki końcowe analizy ekonomicznej składają się miary wyników klinicznych, które opisują i oceniają jakość życia oraz koszty leczenia. Są to wyniki, które stanowią podstawę oceny ekonomicznej (17).

Kolejny dokument zawiera holenderskie procedury oceny leków refundowanych w ramach ambulatoriów. Raport został przygotowany przez Ministerstwo Zdrowia, Opieki Społecznej i Sportu oraz Kolegium Ubezpieczeń Zdrowotnych (College voor Zorgverzekeringen, CVZ).

Podmioty odpowiedzialne dla refundowanych leków mogą formalnie zwrócić się do CVZ o poradę, przed złożeniem formalnego wniosku, o włączenie leku do systemu refundacji leków (Geneesmiddelenvergoedingssysteem, GVS). Dzięki opinii naukowej CVZ można ujawnić możliwe wady i braki wniosku w stosunku do wymaganych standardów, w tym wyboru klinicznie istotnych punktów końcowych. Ponadto opinia CVZ daje możliwość identyfikacji i omówienia ewentualnych problemów metodologicznych w stosowaniu wytycznych farmakoekonomicznych.

Oprócz porównania ocenianego leku z odpowiednim komparatorem lub standardowym leczeniem, ważne jest, aby wybrać odpowiednie punkty końcowe, które są niezbędne do określenia wartości terapeutycznej leku. Wnioskodawca powinien wskazać, w oparciu, o który punkt końcowy, opiera się CEA. Jeżeli we wniosku uwzględniono analizę farmakoekonomiczną, użyte stany zdrowia powinny być spójne z punktami końcowymi, wykorzystywanymi do określania wartości terapeutycznej leku.

Wartość terapeutyczna leku jest określana przez równowagę między zamierzonymi i niezamierzonymi efektami zdrowotnymi ocenianego leku, w stosunku do wybranego komparatora lub standardowego leczenia. Twierdzenie o efektywności terapeutycznej leku jest ograniczone, gdy dostępnych jest niewiele danych na temat klinicznie istotnych punktów końcowych. W przypadku, gdy odnotowano równowagę efektów zdrowotnych, tj. korzyść nie przewyższa ryzyka stosowania, znaczącą rolę mogą odgrywać drugorzędne kryteria oceny (możliwość zastosowania, doświadczenie i łatwość użycia), w zakresie, w jakim stają się widoczne poprzez zamierzone (skuteczność) i/lub niezamierzone (działania niepożądane) efekty. Koszty nie odgrywają żadnej roli w określaniu wartości terapeutycznej.

Złotym standardem określania zamierzonych efektów leczenia jest randomizowane, podwójnie zaślepienie badanie porównawcze. Zamierzone efekty powinny być wyrażone klinicznie istotnymi punktami końcowymi, które są zauważalne dla pacjenta, np.: wskaźnik zachorowalności, śmiertelność i/lub jakość życia.

Klinicznie istotne punkty końcowe często są niedostępne w momencie oceny (np. leki stosowane w celu zapobiegania zdarzeń sercowo-naczyniowych). Takie badania kliniczne zwykle umożliwiają uzyskanie zastępczego wyniku (znanego również jako wynik pośredni). W takich przypadkach zastępcze punkty końcowe (np. wynik laboratoryjny lub charakterystyka fizyczna) tworzą wyłączne parametry użyteczne do oceny zamierzonych efektów. Należy wspomnieć, że powinien istnieć wyraźny związek między punktem zastępczym, a parametrem klinicznym. Surogaty nie zawsze są odczuwalne przez pacjentów. W celu ustalenia odpowiednich miar wyników, Komitet ds. Refundacji Produktów Leczniczych (Commissie Farmaceutische Hulp, CFH) może korzystać z wytycznych EMA oraz wytycznych klinicznych.

Niezamierzony efekt kliniczny to efekt, który wystąpił u pacjenta, pomimo zastosowania prawidłowej dawki leku w celu zapobiegania, diagnozowania lub leczenia choroby. Większość niezamierzonych efektów to negatywne, niezamierzone działania niepożądane leku, chociaż za takie można uważać również wystąpienie odporności bakterii z powodu stosowania antybiotyków. Niezamierzone efekty zdrowotne mogą różnić się dotkliwością, częstotliwością i znaczeniem klinicznym. Im mniej poważne zaburzenia, tym mniejsza akceptowalność zdarzeń niepożądanych.

Porównując różne niezamierzone efekty, szczególną uwagę należy zwrócić na poważne zdarzenia niepożądane i zdarzenia niepożądane występujące z najwyższą częstotliwością. Poważne zdarzenie jest definiowane jako efekt uboczny, który prowadzi do śmiertelności, sytuacji zagrożenia życia, inwalidztwa lub niepełnosprawności, hospitalizacji lub wydłużenia okresu hospitalizacji. Nieoczekiwane zdarzenie niepożądane jest skutkiem ubocznym, które nie zostało opisane w oficjalnym dokumencie rejestracyjnym. Ryzyko wystąpienia niespodziewanego niekorzystnego zdarzenia zmniejsza się, gdy zwiększa się doświadczenie świadczeniodawców z lekiem. Dlatego stwierdzenie dotyczące bezpieczeństwa leku, zawsze powinno uwzględniać aspekt doświadczenia w jego stosowaniu (18).

### 3.9 Irlandia

Wytyczne Health Information and Quality Authority (HIQA) z 2018 roku dotyczą oceny efektywności klinicznej technologii medycznych w Irlandii.

Punkty końcowe muszą być jasno zdefiniowane i mierzalne oraz być wiarygodne i ważne. Powinny być istotne dla leczonej choroby i wrażliwe na zmianę. Wybór punktów końcowych zależy od celu badania lub uwzględnionego porównania. Jeżeli głównym celem stosowania technologii jest poprawa życia, śmiertelność będzie właściwym punktem końcowym. Jeżeli technologia jest zaprojektowana do poprawy mobilności, wtedy status funkcjonalny może być bardziej odpowiednim punktem końcowym. W ocenie wyboru właściwych punktów końcowych pomocni mogą być pacjenci i klinicyści.

HIQA wyodrębnia następujące punkty końcowe: kliniczne punkty końcowe (których wybór musi być uzasadniony na podstawie jasnego związku pomiędzy procesem choroby, technologią i punktem końcowym), punkty końcowe raportowane przez pacjenta (ang. patient-reported outcome, PROs), surogaty, złożone punkty końcowe, zdarzenia niepożądane.

Kliniczny punkt końcowy jest aspektem klinicznym pacjenta lub stanu zdrowia, który mierzy się w celu oceny skuteczności lub szkodliwości leczenia w stosunku do najlepszej, dostępnej alternatywy. Powinien być prawidłowym pomiarem klinicznej korzyści wynikającej z leczenia – to jest klinicznie istotnym, wrażliwym (odpowiadającym na zmiany) oraz rozpoznawanym, jak i stosowanym przez lekarzy. Kliniczne punkty końcowe ustala się na podstawie obecności lub braku klinicznych zdarzeń. Mają one tendencję do bycia niedwuznacznymi, bezstronnymi pomiarami zdarzeń, które stosuje się, żeby zminimalizować potencjalne błędy. Przykładem istotnych klinicznie punktów końcowych są: śmiertelność, udar, amputacja kończyny dolnej.

**Mocne strony:** kliniczne punkty końcowe wydają się być aktualne, jak i wiarygodne. Są zazwyczaj obiektywnymi miarami, redukującymi występowanie błędów oceny, a także uogólniane wśród zestawień i dlatego mogą poprawiać zewnętrzną wiarygodność badania.

**Ograniczenia:** mogą być słabo zdefiniowane w badaniach (np. czy zawał mięśnia sercowego niezakończony zgonem obejmuje ciche zdarzenia „silent events”?). Rozbieżności w definicji mogą prowadzić do różnic w wynikach. Mogą być rzadkimi zdarzeniami, które obniżają siłę statystyczną i wymagają wykonywania prób na dużych populacjach. Niektóre kliniczne punkty końcowe mogą być jednocześnie klinicznie istotne, ale o małym, bezpośrednim znaczeniu dla pacjenta.

PROs powinny być stosowane do mierzenia zmian w zdrowiu i stanie funkcjonalnym, które są bezpośrednio związane z pacjentem. Zazwyczaj odnoszą się do samodzielnego raportowania stanu zdrowia przez pacjentów, skupiając się na tym jak pacjent funkcjonuje lub jakie ma odczucia w związku z chorobą i jej leczeniem. Obejmują proste pomiary objawów (np. ból mierzony skalą Likert), bardziej skomplikowane (np. codzienna aktywność), wielowymiarowe pomiary (np. jakość życia) i satysfakcję z leczenia. PROs powinny być wrażliwe na zmiany w stanie zdrowia. PROs mogą być generyczne (stosowane w jakiegokolwiek chorobie, ale mogą być mniej wrażliwe na zmiany w stanie zdrowia niż pomiary specyficzne choroby) lub specyficzne dla choroby. Jeżeli stosuje się miary wielowymiarowe powinno się jasno stwierdzić czy wszystkie wymiary były oceniane. PROs powinny zawierać domeny istotne dla leczonej choroby. Pomiar jakości życia może być podatny na zmiany, z powodu różnych czynników zewnętrznych (takich jak: okoliczności niezwiązane z leczoną chorobą), z wyjątkiem kwestionariuszy jakości życia, które są specyficznym rozwijane, aby uwzględnić wpływ na proces specyficznej choroby. Możliwe jest, żeby zmapować pomiar jednej miary jakości życia na inny, np. EQ-5D na SF-36. Jednak ten proces musi być jasno określony i uzasadniony. Powinny być stosowane tylko zwalidowane funkcje mapowania na podstawie empirycznych danych. Właściwości statystyczne funkcji mapowania powinny być jasno opisane. Wszystkie PROs zebrane w badaniu powinny być raportowane.

**Mocne strony:** PROs mierzą zmiany w zdrowiu i stanie funkcjonalnym, które mają bezpośrednie znaczenie dla pacjenta. Mogą także wskazywać, gdzie klinicyści i pacjenci mają rozbieżne punkty widzenia na temat punktów końcowych, uważanych za ważne dla pacjenta. PROs mogą obejmować zarówno pozytywne, jak i negatywne efekty technologii w jednym pomiarze.

PROs mogą być stosowane do wykrycia punktów końcowych, takich jak: ból, który może być trudny lub niewykonalny w pomiarze w klinicznym teście.

PROs są często zbierane w formie samodzielnego kwestionariusza, który nie musi być wypełniony w warunkach klinicznych. Nie ma wymogów dla biologicznych próbek i mogą być one oceniane przez nie-klinicystów.

**Ograniczenia:** niektóre generyczne miary PROs wykazały, że są niewrażliwe na umiarkowane zmiany w stanie zdrowia. Jeżeli PROs nie są wrażliwe na zmianę, mogą nie być zdolne do właściwego uchwycenia efektu technologii.

Klinicznie istotne PROs mogą być trudne do określenia, z wyjątkiem przypadku, w którym PRO jest głównym punktem końcowym efektywności leczenia, np. ból stosowany do oceny skuteczności leków przeciwbólowych.

Wyniki PROs są często niespecyficzne dla konkretnego stanu i są często podatne na ogólne zmiany okoliczności, w których znajduje się pacjent, co utrudnia bezpośrednie powiązanie zmiany wyniku z oceną technologii.

PROs często zwracają wynik, co może być trudne do przełożenia zmiany w wyniku na marker poprawy klinicznej. Pojęcia takie jak „minimalna dostrzegalna poprawa kliniczna” i „odpowiedź” są stosowane do określenia klinicznie istotnej poprawy. Definicje klinicznie istotnych zmian i „odpowiadających” są otwarte na pytania i muszą być jasno uzasadnione, jeżeli są stosowane.

Kwestionariusze PROs mogą być czasochłonne do wypełnienia. Jeżeli kohorta pacjentów ma problem z pisaniem i czytaniem, wtedy wskaźnik odpowiedzi może być niższy lub odpowiedzi mogą nie odzwierciedlać dokładnie prawdziwego postrzegania pacjentów. Dane PROs zbierane są wśród pacjentów, którzy nie są zaślepieni podczas przydziału leczenia, co może być stronicze. Jeżeli istnieje taka możliwość, podczas kompletowania pomiarów PROs, aby uniknąć błędów, uczestnicy badania powinni, być zaślepieni w stosunku do przydzielonej terapii.

Surogatowe punkty końcowe muszą mieć jasne uzasadnienie biologiczne lub medyczne albo mieć silny i zwalidowany związek z rozważanym, ostatecznym punktem końcowym.

**Mocne strony:** gdy istnieje wyraźny i silny związek z ostatecznym punktem końcowym, surogat może umożliwić skrócenie okresu obserwacji i znacznie zredukować koszty badania.

**Ograniczenia:** jeżeli mechanizm działania technologii nie jest w pełni zrozumiały, możliwe jest, że za pomocą surogatów nie będzie można dokładnie przewidzieć prawdziwego efektu klinicznego technologii. Co więcej, jeżeli istnieje wiele ścieżek przyczynowych pomiędzy technologią a klinicznym punktem końcowym, wtedy za pomocą surogatów także nie można dokładnie przewidzieć efektu klinicznego. Wielkość efektu na surogat może być znacząco różna od tej na ostateczny punkt końcowy. Dlatego zastosowanie surogatu może prowadzić do niedoszacowania lub przeszacowania efektu technologii. W zależności od ścieżki leczenia surogat może nie uchwycić efektu kolejnych technologii na ostateczny punkt końcowy. To może być szczególnie istotne w złożonych interwencjach lub badaniach naprzemiennych (cross-over), gdzie sekwencje leczenia mogą mieć ważny wpływ na wyniki.

Metody statystyczne stosowane do przewidywania wyników ostatecznych z surogatów nie są doskonałe i nie negują w żaden sposób potrzeby zbierania danych dla ostatecznych punktów końcowych. Analizy przewidujące wpływ na ostateczne punkty końcowe z surogatów powinny być podtrzymywane przez rozległe analizy wrażliwości.

Złożone punkty końcowe – zmiana w tych punktach powinna być klinicznie istotna. Wszystkie indywidualne składowe tego punktu muszą być wiarygodnymi i ważnymi punktami końcowymi. Komponenty, które wpływają na złożony punkt końcowy powinny być zidentyfikowane i podkreślone w analizie.

Wszystkie zdarzenia niepożądane, które są klinicznie lub ekonomicznie ważne muszą być raportowane. Zarówno ciężkość, jak i częstość szkód powinny być zanotowane. Powinno być jasne czy szkoda jest krótkoterminowa czy trwała (19).

Wytyczne HIQA z 2018 roku dotyczą oceny ekonomicznej technologii medycznej w Irlandii.

W analizie ekonomicznej preferowanym typem jest CUA z punktami końcowymi wyrażonymi w QALY. W wyjątkowych sytuacjach można zastosować CEA z wynikami zdrowotnymi wyrażonymi w jednostkach naturalnych, np. LYG (lub innych właściwych punktach końcowych, jeżeli technologia nie wpływa na przeżycie). CEA może być wtórną analizą, gdy możliwe jest zastosowanie punktów końcowych ważnych dla pacjenta (innych niż QALY).

QALY jest najczęściej stosowanym punktem końcowym w CUA. Może jednocześnie łączyć zmiany ilości i jakości życia, z wyższością jednej technologii nad drugą, które wyrażono w zyskanych QALY.

Jeżeli korzyść w CEA mierzy się surogatem lub pośrednim punktem końcowym musi istnieć dobrze ustalone, zwalidowane połączenie między tym markerem a ważnym dla pacjenta punktem końcowym. Powinno się uzasadnić ekstrapolację zmiany surogatowych markerów na klinicznie istotne efekty.

CMA wykonuje się, gdy punkty końcowe (dotyczące skuteczności i bezpieczeństwa) są lub mogą być identyczne (20).

### 3.10 Kanada

Dokument pan-Canadian Oncology Drug Review (pCODR) z 2019 roku dotyczy procedur przeprowadzania przeglądów i wydawania onkologicznych rekomendacji refundacyjnych.

We wnioskach refundacyjnych jednym z priorytetowych kryteriów przeglądu dla leczenia nowotworów jest zastosowanie potencjalnie istotnej poprawy w istotnych punktach końcowych, takich jak (ale nie jest do nich ograniczone):

- poprawa przeżycia całkowitego w adjuwantowych wskazaniach,
- eliminacja lub istotna redukcja działań niepożądanych związanych z leczeniem standardowym,
- mierzalna i istotna poprawa jakości życia nad innymi, dostępnymi terapiami w Kanadzie.

Zespół ds. Przeglądu podczas opracowywania protokołu rozważa punkty końcowe istotne dla pacjenta i kwestie zidentyfikowane przez grupę pacjentów (lub indywidualnego pacjenta/opiekuna, gdy brak grupy pacjentów) (21).

Wytyczne Canadian Agency for Drugs and Technologies in Health (CADTH) z 2017 roku dotyczą oceny ekonomicznej technologii medycznych.

W analizie ekonomicznej referencyjna powinna być CUA, wyrażona w QALY. W CEA wyniki wyraża się w jednostkach naturalnych, takich jak: LYG, uratowane życie lub uniknięcie/osiągnięcie klinicznego zdarzenia.

Analiza kosztów konsekwencji powinna być komplementarna, ale niezastępująca CUA. W analizie tej wyniki są prezentowane jako koszty i wyniki zdrowotne (np. przewidywane zdarzenia, OS, zyskana jakość życia).

CMA przeprowadza się, gdy wszystkie ważne punkty końcowe są równoważne między porównywanymi interwencjami.

Dobłą praktyką jest uwzględnianie najodpowiedniejszych punktów końcowych dla każdego interesariusza.

W analizie ekonomicznej wyróżnia się następujące typy punktów końcowych:

- zdrowotny punkt końcowy, istotny dla pacjenta i nieformalnych opiekunów, np. jakość życia, LYG, kliniczna zachorowalność;
- niezdrowotny punkt końcowy, istotny dla pacjenta i nieformalnych opiekunów, np. informacje dostępne dla pacjenta, redukcja zachowań kryminalnych, osiągnięcie lepszej edukacji.

Decydenci biorą pod uwagę wpływ interwencji na pacjenta leczonego w rutynowej praktyce. W wariantcie referencyjnym wymagałoby to uwzględnienia klinicznie istotnych punktów końcowych, informujących o trwaniu i jakości życia. Interwencja zdrowotna powinna być oceniona także pod względem zagrożeń (zdarzeń niepożądanych). Ważne jest, aby rozważyć zdarzenia niepożądane związane z interwencjami ocenianymi w analizie ekonomicznej, ponieważ mogą one wpływać na spodziewane koszty i wyniki. W szczególności zdarzenia niepożądane w zależności od ich charakteru, częstości, czasu trwania i ciężkości mogą wpływać na przestrzeganie zaleceń przez pacjentów, śmiertelność, zachorowanie, preferencje (użyteczność) lub wykorzystywane zasoby. Badacz powinien skupić się na zagrożeniach, które są klinicznie istotne i w związku z tym mogą mieć wpływ na spodziewane koszty i efekty porównywanych interwencji w praktyce klinicznej.

Surogat jest stosowany w badaniach klinicznych jako zamiennik dla bezpośrednich miar preferencji pacjenta, funkcjonowania i przeżycia. W przypadku braku interesujących parametrów surogaty wskazują skuteczność lub szkodę związaną z porównywanymi interwencjami.

Badacze powinni ocenić i uzasadnić ważność jakiegokolwiek surogatowego punktu końcowego stosowanego w celu oszacowania parametru. Niepewność surogatu w odniesieniu do twardego, klinicznego punktu końcowego powinna być odzwierciedlona w referencyjnym wariantcie analizy probabilistycznej. Istnienie wielu potencjalnych surogatów powinno być odzwierciedlone w analizie niepewności. Kiedy rozważa się użycie biomarkera, jako surogatu, badacz powinien ocenić i uzasadnić ważność biomarkera i stopień, w którym spełnia on kryteria surogatowego punktu końcowego (22).

### 3.11 Niemcy

Raport Instytutu Jakości i Efektywności w Opiece Zdrowotnej (Institute for Quality and Efficiency in Health Care, IQWiG) z 2017 roku dotyczy generalnych metod prawnych i naukowych.

Pytanie kliniczne zawiera m.in. kryteria odnoszące się do punktów końcowych (np. istotnych dla pacjenta). Korzyść interwencji powinna być związana z pacjentem, w związku z tym ocenę powinno się wykonywać na podstawie wyników z badań oceniających wpływ interwencji na punkty końcowe, istotne dla pacjenta.

Punkty końcowe, istotne dla pacjenta odnoszą się do tego, jak pacjent się czuje, funkcjonuje i żyje. Należy rozważyć zamierzone, jak i niezamierzone efekty interwencji, które w szczególności pozwalają na ocenę wpływu na następujące istotne dla pacjenta punkty końcowe (aby określić zmianę związaną z chorobą i leczeniem): śmiertelność, zachorowanie (objawy i komplikacje), jakość życia. Te punkty końcowe powinny być pierwotnie rozważanymi punktami końcowymi.

Powinno się właściwie rozważyć punkty końcowe związane z korzyścią dla pacjenta: wzrost spodziewanej długości życia, poprawę w jakości życia, skrócenie choroby i działania niepożądane.

Działania niepożądane odnoszą się do wszystkich zdarzeń i działań, reprezentujących indywidualne postrzeganie lub obiektywnie, wykrywalne, fizyczne lub psychiczne szkody, które mogą w większym lub mniejszym stopniu spowodować krótkoterminową lub długoterminową redukcję w spodziewanej długości życia, wzroście zachorowalności lub pogorszeniu jakości życia. Trzeba zanotować, że termin „działania niepożądane” jest stosowany wtedy, gdy przypuszcza się związek przyczynowy z interwencją, mając na uwadze, że kwestia związku przyczynowego pozostaje otwarta dla terminu „zdarzenia niepożądane”.

W specjalnych aspektach oceny korzyści wymieniono: punkty końcowe raportowane przez pacjenta, złożone punkty końcowe i subiektywne punkty końcowe.

Punkty końcowe raportowane przez pacjenta mogą zawierać się w odpowiednich wymiarach korzyści dla pacjenta. Oprócz jakości życia, PROs mogą zawierać inne wymiary korzyści, np. objawy choroby, satysfakcję z leczenia.

Złożone punkty końcowe to grupa zdarzeń określona przez badacza (np. zawał mięśnia sercowego, udar i zgon). W tym kontekście indywidualne zdarzenia w tej grupie często różnią się ciężkością i znaczeniem dla pacjenta i lekarza (np. przyjęcie do szpitala i zgon z powodu sercowo-naczyniowego). Warunkiem do rozważenia złożonego punktu końcowego jest to, że jego poszczególne komponenty reprezentują punkty końcowe, istotne dla pacjenta, które są określone w planie raportu. W tym przypadku surogaty mogą być zaakceptowane tylko, gdy Instytut uznał je za ważne. Wyniki poszczególnych komponentów także powinny być raportowane oddzielnie.

Subiektywne punkty końcowe są badane w niezaślepienych RCT np. PROs lub punkty, których dokumentacja i ocena silnie zależy od osób leczących lub oceniających wyniki. Empiryczne badania pokazują, że efekty subiektywnych punktów końcowych są zaburzone w średniej, faworyzując testowaną interwencję. W tej sytuacji Instytut zazwyczaj nie wyciąga wniosków dotyczących korzyści czy szkody z wyników istotnych statystycznie.

Ważne surogaty mogą być rozważone w ocenie korzyści. Surogaty są często stosowane w badaniach medycznych jako zamiennik dla punktów końcowych, istotnych dla pacjenta, głównie w celu wyciągnięcia wcześniejszych i łatwiejszych wniosków o korzyściach (wartości dodanej) istotnych dla pacjenta. Większość surogatów jest jednak niewiarygodna w tym względzie i może wprowadzać w błąd, kiedy są wykorzystywane w ocenie korzyści. W związku z czym surogaty są zazwyczaj stosowane w ocenie korzyści przez Instytut, tylko jeżeli zostały zwalidowane wcześniej za pomocą właściwych metod statystycznych w wystarczająco ograniczonej populacji pacjentów i z porównawczą interwencją (np. lek z porównywalnym sposobem działania). Surogat może być uznany za ważny, jeżeli efekt interwencji na punkt końcowy istotny dla pacjenta jest w wystarczającym stopniu wytłumaczony przez efekt surogatu. Konieczność oceny surogatów może mieć szczególne znaczenie w zakresie wczesnej oceny korzyści leków, ponieważ procedury zatwierdzające pierwotnie badają skuteczność leku, ale nie zawsze korzyść istotną dla pacjenta lub dodatkową korzyść.

Surogaty, które nie są ważne lub dla których brak jest właściwie zwalidowanej procedury mogą być zaprezentowane w raporcie Instytutu. Jednak niezależnie od obserwowanych efektów, takie punkty końcowe nie są odpowiednie do dostarczenia dowodów dla weryfikacji korzyści interwencji.

Wyniki istotne dla pacjenta mogą zostać przekształcone w pomiar całkowitej korzyści, np. wyrażonej w QALY czy DALY (ang. disability-adjusted life year) (23).

### 3.12 Norwegia

Raport Wydziału Medycznego, Norweskiego Instytutu Zdrowia Publicznego z 2017 roku dotyczy oceny przeglądów systematycznych i analiz HTA publikowanych przez ww. Instytucję, dawniej zwaną Norweskim Centrum Wiedzy dla Służby Zdrowia (ang. Norwegian Knowledge Centre for the Health Services) w 2015 roku.

Poniżej zestawiono wybrane wnioski Instytutu z oceny ww. przeglądów systematycznych i analiz HTA:

- Istnieje potrzeba przeprowadzenia badań bezpośrednio porównujących odpowiednie interwencje, badań, które prezentują istotne informacje i punkty końcowe z dłuższych okresów obserwacji.

- W odniesieniu do stosowania naltreksonu w niskich dawkach poza wskazaniem rejestracyjnym najważniejsze punkty końcowe z punktu widzenia pacjenta to zmiany w przebiegu choroby i wpływ na codzienne funkcjonowanie.
- W badaniach dotyczących oceny organizacji opieki zdrowotnej wśród istotnych punktów końcowych wyróżnia się: czas oczekiwania na leczenie, liczba wizyt, zadowolenie pacjentów, częstość korzystania z usług medycznych w nagłych wypadkach, dotrzymywanie terminów świadczeń i ciągłość opieki.
- Nowe badania dotyczące skuteczności szkoleń z zakresu pierwszej pomocy powinny uwzględniać następujące punkty końcowe: podtrzymanie umiejętności ratowania w czasie, przeżycie, punkty końcowe skierowane do osób otrzymujących pierwszą pomoc, współczynnik udzielenia pomocy, czas odpowiedzi poszkodowanego na pomoc oraz jakość świadczonej pierwszej pomocy (24).

### 3.13 Nowa Zelandia

Wytyczne Pharmaceutical Management Agency (PHARMAC) z 2017 roku dotyczą wniosków refundacyjnych (25) a dokument PHARMAC z 2015 roku odnosi się do postępowania w analizach farmakoekonomicznych (26).

PHARMAC zaleca, aby w składanych wnioskach przedstawiać wyniki dla ,m.in.: pierwszorzędowych punktów końcowych, prospektywnie zdefiniowanych, drugorzędowych punktów końcowych i działań niepożądanych w populacji zgodnej z intencją leczenia (ITT) (25,26). Punkty końcowe pierwszorzędowe i drugorzędowe powinny obejmować wyniki publikowane, jak i niepublikowane (25).

W analizie ekonomicznej zaleca się wykorzystywanie statystycznie i klinicznie istotnych punktów końcowych. Preferowane są te pochodzące z wysokiej jakości randomizowanych badań klinicznych (RCT), przeglądów systematycznych lub metaanaliz (stopień dowodów 1+ lub 1++) (25,26).

Analiza ekonomiczna powinna mieć formę CUA z korzyściami mierzonymi w QALY (25,26). QALY łączy zmiany w ilości i jakości życia (śmiertelności i zachorowalności) w jedną, złożoną miarę (26). W CEA uwzględnia się jasno zdefiniowane punkty końcowe (wyrażone w jednostkach fizycznych), np. LYG, uniknięte hospitalizacje, zapobiegnięcie udarowi lub zwichnięciu biodra (25,26). W przypadku, gdy między ocenianym leczeniem a komparatorem wykazano równoważność w klinicznych punktach końcowych właściwą analizą jest CMA (25,26).

Ważne jest, aby upewnić się, że do CUA włączono punkty końcowe najodpowiedniejsze dla choroby, które odzwierciedlają perspektywę i zakres modelu. Analiza ekonomiczna powinna być wykonana w oparciu o raportowane, klinicznie istotne punkty końcowe, ważne dla zdrowia pacjenta, które wpływają na progresję choroby, przeżycie całkowite i jakość życia. Czasem mogą być dostępne tylko surogaty. Są one zamiennikami dla klinicznie istotnych punktów końcowych i pozwalają zmierzyć jak pacjent się czuje, funkcjonuje lub żyje. Surogaty mogą być stosowane w CUA, tylko gdy nie ma dostępnych alternatywnych danych. Z ostrożnością powinno się je stosować, gdy nie mają przełożenia na klinicznie istotne punkty końcowe i wyniki efektywności. Jeżeli istnieje niepewność dotycząca klinicznej istotności punktów końcowych lub korelacji między surogatem a klinicznymi punktami końcowymi powinno się stosować konserwatywne założenia dotyczące przeżycia całkowitego i jakości życia (26).

### 3.14 Szkocja

Wytyczne szkockiej agencji HTA (Scottish Medicine Consortium, SMC) dla wnioskodawców dotyczące wypełnienia formularza oceny nowego produktu (ang. New Product Assessment Form).

W części dotyczącej skuteczności należy przedstawić kliniczne dowody naukowe na stosowanie leku w proponowanej populacji. Dane powinny pochodzić z kluczowych/głównych badań dla pierwszorzędowych lub drugorzędowych punktów końcowych dla analizowanej podgrupy pacjentów.

Dla pierwszorzędowego punktu końcowego należy określić jego definicję, w tym szczegóły dotyczące metod gromadzenia danych i czas oceny. Jeśli pierwszorzędowy punkt końcowy jest mierzony w systemie punktacji, należy podać krótkie szczegóły, w tym wskazanie trafności wyniku (np. wyższe wyniki = lepsza jakość życia).

Dla każdej wtórnej analizy pierwszorzędowego punktu końcowego (np. analiza w podgrupie, w której lek jest zarejestrowany) lub analizy istotnych drugorzędowych punktów (np. przeżycie, gdy podstawowym wynikiem była odpowiedź guza) należy podać szczegółową charakterystykę badanej populacji uwzględnionej w tych analizach, metody uwzględnienia brakujących danych i szczegóły zastosowanych testów statystycznych. Jeśli którakolwiek z tych analiz została zaprojektowana post-hoc, należy podać szczegóły uzasadniające przeprowadzenie analizy post-hoc.



Preferowana jest subanaliza pierwszorzędowych lub drugorzędowych punktów końcowych badań aktywnie kontrolowanych. Dopuszczane są badania kontrolowane placebo i niekontrolowane, jeżeli dostarczają odpowiednich dowodów.

Należy podać szczegóły, czy w badaniach bezpośrednio mierzono wyniki zdrowotne, takie jak: śmiertelność, przeżycie, częstość występowania choroby, zachorowalność, sprawność funkcjonalna, jakość życia lub czy mierzono surogaty: obniżenie ciśnienia krwi. Należy podać szczegóły dotyczące jakiegokolwiek związku między surogatami, a korzyścią zdrowotną lub jej brakiem dla pacjentów.

Należy uwzględnić główne skutki uboczne, które mogą wystąpić.

Oprócz informacji na temat skuteczności klinicznej i bezpieczeństwa, wnioskodawcy powinni rozważyć:

- wyniki badań obserwacyjnych, w tym rejestrów chorób, które mogłyby uzupełniać konwencjonalne dane z badań klinicznych;
- efektywność leku zgłaszaną przez pacjenta w postaci takich wyników, jak: jakość życia związana ze zdrowiem, stan zdrowia, funkcjonowanie fizyczne, aktywność w życiu codziennym, przestrzeganie leczenia, zadowolenie pacjenta z leczenia itp. (27-29).

### 3.15 Szwecja

Dokument wydany przez szwedzką agencję HTA (Swedish Agency on Health Technology Assessment and Assessment of Social Services, SBU) dotyczy oceny metodologii stosowanej w opiece zdrowotnej.

Jednym z elementów pytania klinicznego zgodnego ze schematem PICO są punkty końcowe. Punkty końcowe o bezpośrednim znaczeniu dla pacjenta to np. przeżycie, jakość życia, zachorowalność i zmiany objawów. Efekty zdrowotne mogą obejmować również powikłania i niekorzystne skutki interwencji. Badania nad ekonomiką zdrowia często są wyrażone jako koszt za QALY.

Według SBU miary wyniku zdrowotnego powinny być odpowiednie do określenia efektu interwencji. Wyniki zdrowotne powinny bezpośrednio odzwierciedlać stan zdrowia pacjenta, np. śmiertelność, zachorowalność, cierpienie, upośledzenie funkcjonowania i jakość życia. Ponadto można wybrać zastępcze punkty końcowe, tj. mierzalne czynniki, które w pewien sposób odnoszą się do wyników istotnych dla pacjenta; przykładami są poziom lipoproteiny w surowicy krwi, ciśnienie krwi i gęstość kości.

Powszechne w badaniach klinicznych są również złożone punkty końcowe. Uzasadnieniem łączenia różnych miar wyników jest zwiększenie mocy statystycznej badań. Jednocześnie należy zachować ostrożność przy pomiarach złożonych punktów końcowych, szczególnie w przypadku uwzględnienia zastępczych punktów końcowych. Istotny statystycznie wpływ złożonego punktu końcowego można często tłumaczyć wyłącznie efektami zastępczego punktu końcowego lub zmiennej, która jest mniej istotna dla pacjenta. Możliwe jest również, że złożone punkty końcowe mogą maskować negatywny wpływ leczenia na istotne wyniki, np. zgon i zdarzenia sercowo-naczyniowe.

Oczekiwania dotyczące efektów leczenia mogą również wpływać na pomiar wyniku. Jest to szczególnie ważne dla takich wyników jak jakość życia lub oszacowanie ciężkości objawów. Wyniki takie jak przeżycie lub złamania są mniej narażone na wpływ pacjenta/badacza.

Podając główne wyniki badań, autorzy powinni opierać się na istotnych klinicznie punktach końcowych. Jeżeli nie uzyskano wyniku istotnego statystycznie dla pierwszorzędowego punktu końcowego, autorzy badania mogą posłużyć się punktem końcowym, który okazał się istotny statystycznie. W celu odnalezienia istotnych wyników, dopuszczane są różne typy analiz ad-hoc w podgrupach (30).

### 3.16 Węgry

Praca z 2017 roku opublikowana przez Narodowy Instytut Farmacji i Żywności (National Institute of Pharmacy and Nutrition) stanowi wytyczne na temat metodologii oceny technologii medycznych.

W najwyższym stopniu, w jakim to możliwe ocena technologii medycznych powinna mieć na celu pomiar (długoterminowych) wyników klinicznych osiąganych w rutynowej praktyce, zamiast wyników skuteczności osiąganych w kontrolowanych okolicznościach (korzyść zdrowotna wykazana w ściśle określonych i kontrolowanych badaniach klinicznych).

Trzy domeny klinicznych punktów końcowych to: śmiertelność, przebieg choroby i jakość życia związana ze zdrowiem. Kliniczne punkty końcowe muszą charakteryzować się adekwatnością, czułością, istotnością i odtwarzalnością. W ocenie względnej skuteczności należy ustalić hierarchię punktów końcowych (pierwszorzędowe i drugorzędowe punkty końcowe). Wybór zależy od populacji docelowej, charakterystyki danej

choroby/stanu zdrowia (zagrożający życiu lub niezagrożający życiu) oraz od celu leczenia. W przypadku choroby zagrażającej życiu - śmiertelność, jako punkt końcowy lub punkty końcowe dotyczące przeżycia są pierwszorzędowymi punktami końcowymi, natomiast chorobowość i/lub jakość życia są drugorzędowymi punktami końcowymi. W przypadku chorób niezagrożających życiu preferowane są punkty dotyczące przebiegu choroby i jakości życia. Przy prezentacji wyników należy również odnieść się do znaczenia klinicznego mierzonych zmian.

Zastosowanie zastępczych punktów końcowych (surogatów) jest dopuszczalne, jeżeli twarde punkty końcowe nie są dostępne, a dowody potwierdzają korelację między zastępczymi a twardymi punktami końcowymi.

Zastosowanie niepoddanych walidacji zastępczych punktów końcowych może być konieczne, jeśli nie są dostępne kliniczne punkty końcowe, ani zwalidowane zastępcze punkty końcowe. Należy uzasadnić zastosowanie niepoddanych walidacji zastępczych punktów końcowych.

Warunkiem zastosowania złożonych punktów końcowych jest spełnienie przez każdą składową złożonego punktu końcowego kryteriów adekwatności, odtwarzalności i znaczenia klinicznego (31).

### 3.17 EUnetHTA

Wytyczne EUnetHTA z 2015 roku dotyczą klinicznych punktów końcowych stosowanych w ocenie względnej skuteczności (ang. Relative Effectiveness Assessment, REA).

Wytyczne zawierają zestaw zaleceń dotyczących wyboru i oceny klinicznych punktów końcowych (ang. clinical endpoints) podczas przeprowadzania REA. Wykonanie REA powinno wykazać, czy zastosowane leczenie ma klinicznie i statystycznie istotny wpływ na odpowiedni punkt końcowy w porównaniu z alternatywą w rzeczywistych warunkach.

Kliniczne punkty końcowe są uważane za sposób pomiaru wpływu leczenia na przeżycie, funkcjonowanie i samopoczucie pacjenta. Wpływ ten ma zwykle postać poprawy stanu zdrowia (np. przeżycie, wyleczenie, remisja choroby) lub też pogorszenia stanu zdrowia (np. działania niepożądane, hospitalizacje, zgon) (32).

Podczas przeprowadzania oceny HTA można skupić się na zbadaniu: najpoważniejszych zdarzeń niepożądanych; najczęściej występujących zdarzeń niepożądanych i innych specyficznych zdarzeń niepożądanych istotnych dla klinicystów i pacjentów (33).

Kliniczne punkty końcowe można ogólnie podzielić na trzy domeny: śmiertelność, przebieg/nasilenie choroby i miary jakości życia związanej ze zdrowiem.

Poprawa przeżycia oznacza bezpośrednią korzyść kliniczną dla pacjentów. Typowe wyniki dotyczące przeżycia obejmują: przeżycie całkowite, przeżycie wolne od choroby i przeżycie wolne od progresji choroby. Przeżycie całkowite jest złotym standardem dla wykazania korzyści klinicznych i jako takie powinno być stosowane tam, gdzie to tylko możliwe. Zdefiniowany jako czas od randomizacji do zgonu, jest jednoznaczny i nie podlega interpretacji przez badacza. Oceniając czas przeżycia bez progresji, należy regularnie oceniać pacjentów, aby zapewnić dokładny pomiar czasu wystąpienia progresji.

Punkty końcowe przedkładane do oceny powinny wyraźnie odnosić się do choroby, stanu lub procesu będącego przedmiotem oceny. Kliniczne punkty końcowe powinny być: głównym objawem lub oznaką choroby; ważną miarą korzyści klinicznej spowodowanej leczeniem; klinicznie istotne, wrażliwe (wrażliwe na zmiany) i uznawane/używane przez lekarzy. Kliniczne punkty końcowe powinny być odtwarzalne i prawidłowe. Odtwarzalny punkt końcowy ułatwia porównania między badaniami i jurysdykcjami. Prawidłowy punkt końcowy mierzy to, co zamierzano zmierzyć. Ważność może być obniżona przez ryzyko błędu związanego z selekcją, błędem poznawczym oraz resztkowe czynniki zakłócające. Powinno się wykazać, że zastosowane punkty końcowe są odpowiednie dla celu, w którym mają być zastosowane. Kwestie dotyczące dokładności wyników badań mogą znaleźć odzwierciedlenie w istotności statystycznej efektu leczenia. Klinicznie istotny efekt powinien spełniać standardy lub konsensus co do wielkości i jakości wyniku badania, który jest uważany za znaczący przez niezależnych klinicystów i/lub pacjentów. Precyzyjna definicja, odtwarzalność, ważność oraz znaczenie statystyczne i kliniczne punktu końcowego powinny być wykazane.

Kliniczne punkty końcowe mogą być pośrednie lub ostateczne (ang. final). Pośredni punkt końcowy to kliniczny punkt końcowy związany ze stosowaniem interwencji, taki jak pomiar funkcji lub objawów (przeżycie wolne od choroby, częstość występowania dusznicy bolesnej, tolerancja wysiłku), który oczekuje się, że będzie korelować ze zmianami zaobserwowanymi w ostatecznych punktach końcowych. Nie jest to ostateczny punkt końcowy choroby, taki jak przeżycie lub odsetek osób z nieodwracalnymi zmianami chorobowymi (udar, zawał mięśnia sercowego). Poprawa w pośrednim punkcie końcowym spowodowana leczeniem jest dobrze postrzegana

i może być cenna dla pacjenta, nawet jeśli nie prowadzi do poprawy przebiegu/nasilenia choroby zachorowalności lub śmiertelności.

Surogat to taki punkt końcowy, który ma zastąpić kliniczny punkt końcowy będący przedmiotem zainteresowania w ramach REA. Surogat ma przewidywać korzyść kliniczną lub zagrożenia dla zdrowia na podstawie dowodów epidemiologicznych, patofizjologicznych, terapeutycznych i innych dowodów naukowych. Niektóre przykłady zastępczych punktów końcowych to ciśnienie krwi, jako zastępczy punkt końcowy choroby sercowo-naczyniowej lub miano wirusa HIV1-RNA, jako wskaźnik supresji wirusa w interwencjach stosowanych w zakażeniu HIV.

Poddane walidacji pośrednie punkty końcowe mogą być zastosowane, gdy nie jest możliwe zmierzenie długoterminowych lub ostatecznych punktów końcowych. Należy zachować ostrożność przy ekstrapolacji wyników pośrednich na końcowe, chyba że wyraźnie wykazano biologiczne lub medyczne uzasadnienie albo silny lub potwierdzony związek z ostatecznym punktem końcowym.

Ostateczne punkty końcowe są często określane przez przeżycie i o ile to możliwe, powinny być stosowane zamiast wyników pośrednich. Jeżeli ostateczne punkty końcowe są niedostępne, wyniki pośrednie mogą być akceptowalne, jeśli istnieją przekonujące dowody na wyraźną i spójną korelację z ostatecznym wynikiem.

Ostateczne punkty końcowe odnoszą się do ostatecznego celu terapeutycznego zastosowania interwencji w opiece zdrowotnej, a nie tylko do wyników klinicznych, dlatego mają one większe znaczenie dla pacjenta i dla ogólnej priorytetyzacji. W stosownych przypadkach ostateczne punkty końcowe powinny być zdefiniowane jako wskaźniki przeżycia, które odzwierciedlają prawdopodobieństwo lub częstotliwość przeżycia w ściśle określonym przedziale (np. uzyskane lata życia). Ocena końcowych punktów końcowych umożliwia porównanie różnych interwencji pod warunkiem porównywalności ostatecznych skutków. Należy zauważyć, że związek między surogatem a punktem końcowym stanowiącym przedmiot zainteresowania nigdy nie można uznać za pewny. Związek ten, nawet jeśli jest dobrze ugruntowany dla danej technologii medycznej, może zostać zakwestionowany inną technologią medyczną, która zapewnia taki sam wpływ na surogat, ale nieoczekiwany wpływ na ostateczny punkt końcowy.

Jeżeli jako punkt końcowy zastosowano PFS, powinny istnieć wystarczające niezależne dowody, aby wykazać, że jest to związane z przeżyciem całkowitym.

Złożone punkty końcowe łączą wiele pojedynczych punktów końcowych w jeden punkt końcowy, demonstrując ogólny efekt leczenia. Interpretując złożone punkty końcowe, należy zachować ostrożność, ponieważ wszystkie składowe punkty końcowe powinny spełniać kryteria ważności, odtwarzalności i znaczenia klinicznego.

Znaczenie kliniczne niektórych PROs mogą być trudne do ustalenia, z wyjątkiem przypadków, w których PROs są głównym punktem końcowym skuteczności dla danej choroby (np. ból stosowany do oceny skuteczności środka przeciwbólowego). Zaproponowano koncepcję minimalnej odczuwalnej poprawy klinicznej pacjenta, aby przenieść zmianę wyniku HRQoL na wskaźnik poprawy klinicznej. Jednak znaczenie kliniczne obserwowanej zmiany może być łatwiejsze do interpretacji poprzez definiowanie osób odpowiadających na leczenie (np. osobą odpowiadającą może być pacjent, który zgłasza, że ból zmniejszył się o co najmniej 50% w porównaniu do stanu początkowego), chociaż może to mieć wpływ na zdolność wykrywania korzyści leczenia. Wykazano, że niektóre miary HRQoL nie są dostatecznie wrażliwe, a inne są przesadnie wrażliwe na niewielkie zmiany statusu, dlatego ważne jest ustalenie klinicznie znaczącej różnicy w ocenie.

Powinno być jasne, że zastosowany instrument do oceny HRQoL jest odpowiedni do zastosowania w kontekście użytkowania, wraz z dowodami potwierdzającymi jego ważność, niezawodność i szybkość reakcji, tak aby wyniki można było łatwiej interpretować. Pomiar HRQoL mogą wypuklić sytuacje, w których lekarze i pacjenci mają rozbieżne opinie na temat tego, co uważa się za ważne dla pacjentów. Ponieważ na HRQoL mogą wpływać zarówno korzyści, jak i działania uboczne wynikające z leczenia, a także nie jest to miara bezpośrednio związana z efektem terapii, możliwe jest wykrycie poprawy w pojedynczym klinicznym punkcie końcowym przy braku zmiany HRQoL i odwrotnie. Powinno być jasno ustalone, że zastosowany instrument pomiaru HRQoL jest odpowiedni do zastosowania w danym kontekście, wraz z dowodami potwierdzającymi jego ważność, niezawodność i wrażliwość, tak aby wyniki można było łatwiej interpretować.

Podobnie jak w przypadku każdej miary klinicznego efektu, niezwykle ważne jest, aby miara HRQoL charakteryzowała się istotnością i odtwarzalnością. Istnieją okoliczności, w których lepiej jest zastosować specyficzną dla choroby miarę HRQoL w porównaniu do generycznych miar HRQoL.

Należy zauważyć, że istotny statystycznie wpływ terapii w badaniu, nie zawsze wystarcza, do określenia jego znaczenia klinicznego. W związku z tym nie wystarczy stwierdzić, że wykazano istotny statystycznie efekt lub podać przedział ufności, lub wartość p. Należy określić wielkość efektu i związaną z nim niepewność oraz wyjaśnić jego znaczenie kliniczne (32).

### Surogaty (zastępcze punkty końcowe)

W miarę możliwości REA należy opierać na ostatecznych klinicznych punktach końcowych, istotnych dla pacjenta (np. punkty końcowe odnoszące się do przebiegu/nasilenia choroby, ogólna śmiertelność). W przypadku braku dowodów dotyczących ostatecznych, istotnych dla pacjenta klinicznych punktów końcowych, które bezpośrednio mierzy korzyść kliniczną, zarówno biomarkery, jak i pośrednie punkty końcowe będą uważane za zastępcze punkty końcowe w REA, jeśli mogą w wiarygodny sposób zastąpić kliniczny punkt końcowy i przewidzieć jego korzyść kliniczną.

Jeżeli zastępcze punkty końcowe są stosowane w REA, należy odpowiednio poddać walidacji ich związek z ostatecznymi punktami końcowymi w oparciu o wiarygodność biologiczną i dowody empiryczne. Poziom dowodów dotyczących walidacji, związana z nimi niepewność i ograniczenia związane z ich zastosowaniem należy wyraźnie wyjaśnić. Zawsze należy podać pełne dane dotyczące walidacji. W przypadku odpowiednio zwalidowanych zastępczych punktów końcowych druga walidacja do celów REA nie będzie konieczna.

Walidacja surogatu względem klinicznego punktu końcowego istotnego dla pacjenta jest standardowo przeprowadzana dla konkretnej populacji i dla konkretnej interwencji, tj. walidacja jest zależna od choroby, specyficzna dla populacji i specyficzna dla technologii. Demonstracja walidacji surogatu zarówno w obrębie klas technologii, jak i między nimi, np. klasy leków, powinna być dokładnie uzasadniona.

W przypadku pierwszej oceny, nawet jeśli preferowane są ostateczne punkty końcowe, zastępcze punkty końcowe mogą zostać zaakceptowane, jeżeli ważność związku zastępczego/ostatecznego dla klinicznego punktu końcowego została wcześniej wyraźnie ustalona w odniesieniu do klinicznych punktów końcowych będących przedmiotem zainteresowania REA. Dostępność wystarczająco dużej bazy danych nt. bezpieczeństwa jest szczególnie ważna. Zawsze należy zgłaszać dowody dotyczące wyników w zakresie bezpieczeństwa.

Podczas ponownej oceny należy zasadniczo wykazać, kiedy jest to tylko możliwe, skuteczność związaną z przebiegiem/nasileniem choroby i śmiertelnością (np. udar, zawał mięśnia sercowego, złamanie).

Porównawcze dane kliniczne (lub dowód, że takie dane zostaną dostarczone w rozsądnych ramach czasowych) dotyczące odpowiednich klinicznych punktów końcowych i bezpieczeństwa pochodzących z badań klinicznych po wprowadzeniu do obrotu i innych źródeł, należy w miarę możliwości dostarczyć przed dokonaniem ponownej oceny.

Brak danych dotyczących klinicznych punktów końcowych istotnych dla REA może być akceptowalny, gdy badanie klinicznego punktu końcowego jest trudne lub niemożliwe (bardzo rzadkie występujące lub opóźnione w czasie), lub populacja docelowa jest zbyt mała, aby uzyskać znaczące wyniki w odniesieniu do odpowiednich klinicznych punktów końcowych, nawet po bardzo długim okresie obserwacji (bardzo powoli postępujące i/lub rzadkie choroby). Jednak wyjątki takie należy dokładnie uzasadnić i uzgodnić przed REA.

Wymogi ponownej oceny w odniesieniu do dalszych danych dotyczących istotnych klinicznych punktów końcowych powinny być jasno określone, gdy REA została wcześniej, dla pierwotnej oceny, sporządzona na podstawie zastępczych punktów końcowych.

Akceptowalność zastępczego punktu końcowego jest również oparta na innych względach dotyczących stosunku korzyści do ryzyka i/lub korzyściach dla zdrowia publicznego, takich jak poważna choroba zagrażająca życiu bez alternatywnej terapii, dostępna duża baza danych dotyczących bezpieczeństwa, trudny do zbadania kliniczny punkt końcowy (bardzo rzadki lub opóźniony w czasie).

W dziedzinie onkologii i hematologii, szczególnie gdy istnieje duża niezaspokojona potrzeba i/lub brak jest dostępnej alternatywy terapeutycznej, zastosowanie pośrednich punktów końcowych, takich jak przeżycie bez progresji choroby (PFS), wydaje się dopuszczalne do przyspieszonego/warunkowego dopuszczenia do obrotu. W ramach REA dopuszczalność pośrednich punktów końcowych dla tych interwencji nadal jest różna w różnych krajach UE i jest przeprowadzana indywidualnie dla każdego przypadku. W onkologii PFS jest pośrednim punktem końcowym, który jest istotny sam w sobie. Jednak dane dotyczące tylko PFS bez danych dotyczących przeżycia całkowitego (OS) lub przynajmniej HRQoL lub innych punktów końcowych istotnych dla pacjenta uważa się za niewystarczające dla REA. Ponadto ważny jest moment, w którym dokonuje się REA; decyzje mogą się różnić w zależności od dojrzałości danych na temat końcowych punktów końcowych, na przykład, jeżeli oceny dokonuje się, gdy nie ma wystarczających informacji na temat klinicznie istotnego punktu końcowego lub są one niewystarczające w porównaniu z sytuacjami, w których istnieją dane dotyczące zarówno zastępczego, jak i ostatecznego punktu końcowego. Dopuszczalność przeżycia bez progresji nie ma takiego samego znaczenia w przypadku choroby adjuwantowej lub przerzutowej. W warunkach leczenia adjuwantowego PFS wydaje się akceptowalny; w przypadku choroby przerzutowej sam PFS jest niewystarczający. PFS można rozważyć

## Podjęcie światowych agencji HTA do istotności klinicznej punktów końcowych

w połączeniu z oceną jakości życia i danymi dotyczącymi przeżycia, których dojrzałość zostanie rozpatrzona indywidualnie dla każdego przypadku (34).

### Złożone punkty końcowego (ang. composite endpoint, CE)

Złożone punkty końcowe zasadniczo nie powinny być stosowane, jeśli dostępny jest odpowiedni pojedynczy pierwszorzędowy punkt końcowy. Jeśli pojedynczy pierwszorzędowy punkt końcowy nie jest dostępny lub jeśli złożony punkt końcowy można uzasadnić jako bardziej odpowiedni (np. rzadka choroba/zdarzenie kliniczne), można wybrać złożony punkt zamiast pojedynczego pierwszorzędowego punktu końcowego.

Powinny istnieć wcześniejsze empiryczne i kliniczne dowody nt. wartości każdego wybranego komponentu złożonego punktu końcowego. Liczba składników złożonego punktu końcowego powinna być ograniczona do trzech lub czterech, aby uniknąć problemów w analizie i interpretacji.

Badania wykorzystujące złożone punkty końcowe powinny być przeprowadzane zgodnie z wytycznymi CONSORT i powinno się zgłaszać wcześniej określone pierwszorzędowe i drugorzędne punkty końcowe, aby umożliwić odpowiednią ich interpretację. Zmiany w definicji złożonego punktu końcowego nie powinny wystąpić podczas badania. Wszystkie składniki złożonego punktu końcowego należy osobno zdefiniować jako drugorzędowe punkty końcowe i raportować w wynikach pierwotnej analizy.

Najlepiej łączyć składniki o podobnym znaczeniu klinicznym i wrażliwości na interwencję. Należy unikać heterogeniczności (połączenia subiektywnych i obiektywnych punktów końcowych). Należy także unikać włączenia składników, w przypadku których wiadomo, że wpływ interwencji jest niewielki lub mało prawdopodobny. Jeśli jest to odpowiednie, należy jednak uwzględnić śmiertelność, gdy istnieje prawdopodobieństwo, że wpływa ona na cenzurowanie obserwacji innych składników. Złożone punkty końcowe można wykorzystać do oceny nie tylko skuteczności, ale także negatywnych efektów technologii medycznych.

Efekty leczenia należy w pierwszej kolejności zgłaszać dla CE. Wyniki należy również podać dla każdego składnika złożonego punktu końcowego; w sposób, w jaki przyczynił się on do wyniku ogólnego, złożonego punktu końcowego. Wszystkie wyniki należy zgłaszać osobno, nawet jeśli nie mają one mocy statystycznej. Lista wyników dla wszystkich składników powinna być przedstawiona w tabeli z podaniem odpowiednich przedziałów ufności.

Poszczególne składniki CE można przedstawiać według poziomów dowodów, przykładowo: L1, śmiertelność ogólna; L2, śmiertelność z powodu określonej przyczyny; L3, nieprowadzące do śmierci zdarzenia kliniczne; L4, objawy choroby. W przypadkach, w których złożony punkt końcowy obejmuje zdarzenia śmiertelne i nieprowadzące do zgonu, zaleca się zgłaszanie wyników dotyczących odpowiednich kombinacji składników CE. Wszystkie dane powinny być zgłaszane. Należy szczegółowo podać liczbę pacjentów z częściowo brakującymi wartościami niektórych składników CE. Jeśli istnieją odpowiednie podgrupy lub specjalne populacje pacjentów zwiększonego ryzyka (takie jak osoby w podeszłym wieku lub pacjenci z niewydolnością nerek), należy przedstawić wyniki dla tych podgrup.

Efekty leczenia należy interpretować przede wszystkim na podstawie CE. Jednak należy również przeanalizować wpływ leczenia na każdy składnik złożonego punktu końcowego w sposób, w jaki przyczynił się on do wyniku złożonego punktu końcowego, aby ocenić, czy interwencja miała podobny wpływ na wszystkie składniki CE.

Zaleca się sprawdzenie, czy leczenie nie wpływa negatywnie na ważne klinicznie składniki złożonego punktu końcowego, ponieważ niektóre terapie mogą mieć negatywny wpływ na jeden składnik, który może być maskowany przez duży korzystny wpływ pozostałych składników. W takich przypadkach może nie być możliwe stwierdzenie, że leczenie ma klinicznie istotny wpływ na złożony punkt końcowy jako całość. Proces REA powinien określić i/lub zidentyfikować, który składnik jest głównie odpowiedzialny za ogólny efekt.

Jeśli dostępne są prawidłowe i porównywalne złożone punkty końcowe z kilku badań, należy rozważyć oparcie ogólnego wniosku na wynikach metaanalizy. Jeżeli występuje jeden istotny problem lub znaczna kumulacja problemów związanych z danym CE, należy stwierdzić znaczną niepewność, co do istotności wyników badania. Pozycja tego badania w hierarchii dowodów i jego przydatność dla REA będzie musiała zostać obniżona (35).

### HRQoL

Instrumenty HRQoL stosowane w kontekście REA powinny przede wszystkim być odpowiednie do celów, którym mają służyć. Osoby przeprowadzające REA powinni zatem najpierw rozważyć, w jakim celu zostanie przeprowadzona względna ocena skuteczności: podejmowania decyzji refundacyjnych lub decyzji klinicznych.

Ogólne zalecenie mające zastosowanie do wszystkich rodzajów REA, niezależnie od ich szczególnego celu, stanowi wymaganie włączenia specyficznej dla choroby lub populacji i generycznej metody pomiaru HRQoL określonej choroby lub populacji oraz ogólnego pomiaru HRQoL w celu jak najbardziej odpowiedniego

uchwycenia wpływu choroby na życie codzienne. W przypadku potrzeby oszacowania QALY, zaleca się pomiar użyteczności (metodą handlowania czasem lub metodą loterii) lub zastosowanie generycznego instrumentu pomiaru HRQoL powiązanego z referencyjnym zestawem norm użyteczności.

W przypadku krajów, które wymagają oceny ekonomicznej w celu zatwierdzenia decyzji o refundacji kosztów technologii medycznych, zaleca się stosowanie danych pochodzących z generycznych instrumentów szacowania użyteczności w ramach badań klinicznych. Wartości użyteczności powinny pochodzić od ogółu społeczeństwa (pośredni pomiar użyteczności) lub od pacjentów (bezpośredni pomiar użyteczności). Wymagana jest konsekwencja w stosowaniu wybranego źródła wartości użyteczności.

W krajach, które nie wymagają oceny ekonomicznej w celu zatwierdzenia decyzji o refundacji kosztów technologii medycznych, wystarczające może być zastosowanie specyficznego dla choroby lub generycznego instrumentu mierzenia HRQoL. W tym kontekście przydatne pozostają środki oceny użyteczności dla REA, zwłaszcza do szacowania QALY, które jest szczególnie przydatne do porównywania interwencji wpływających zarówno na śmiertelność, jak i na przebieg/nasilenie choroby.

REA przeprowadzone w celu podejmowania decyzji dotyczących alokacji zasobów pomiędzy różnymi wskazaniami powinno przede wszystkim opierać się na danych HRQoL uzyskanych za pomocą generycznego instrumentu mierzenia HRQoL, obejmującego wszystkie wymiary HRQoL, w których wykazana poprawa jest uważana za istotną przez ogół społeczeństwa.

Jeśli nie obserwuje się poprawy w wynikach uzyskanych za pomocą generycznych instrumentów mierzenia HRQoL, domniemana korzyść z interwencji jest mniej prawdopodobna z punktu widzenia społeczeństwa, biorąc pod uwagę zakres istniejących problemów zdrowotnych, między którymi należy podzielić zasoby publiczne.

REA przeprowadzone w celu informowania o decyzjach dotyczących alokacji zasobów w ramach wskazań może być oparte na zweryfikowanych kompleksowych danych pochodzących ze specyficznych dla choroby instrumentów mierzenia HRQoL, ponieważ porównywalność między wskazaniami jest w tym przypadku mniej istotna.

REA przeprowadzone w celu informowania pracowników służby zdrowia i pacjentów może być oparte na specyficznych dla choroby instrumentach mierzenia HRQoL. Instrumenty mierzenia HRQoL specyficzne dla choroby mogą być przydatne do bardziej szczegółowej oceny ogólnych wymiarów HRQoL, na które wpływ ma oceniana interwencja.

Korzyści w zakresie HRQoL wynikające z zastosowania interwencji należy wykazać za pomocą powtarzanych pomiarów zarówno w grupie interwencyjnej, jak i kontrolnej. Ocena pojedynczego wyniku HRQoL uważa się za niewystarczającą do wykazania względnej skuteczności, ponieważ podlega ona ryzyku błędu systematycznego i często jest zbyt ograniczona, aby wykryć zmiany stanu zdrowia.

Mapowanie instrumentów specyficznych dla choroby lub generycznych instrumentów do instrumentów opartych na preferencjach w celu uzyskania wartości użyteczności zasadniczo nie jest zalecane w ramach REA. Władze powinny zachęcać badaczy, aby zawsze włączali instrument oparty na preferencjach do protokołu badania klinicznego, aby uniknąć konieczności mapowania.

Należy przedstawić dokumentację procesu oceny istotności, wiarygodności, wrażliwości na zmiany i akceptowalności instrumentów oceny HRQoL stosowanych w REA, biorąc pod uwagę zastosowaną metodę zbierania wyników i możliwe adaptacje kulturowe i/lub językowe.

Ocena HRQoL przez pełnomocników/opiekunów (z ang. proxy judge) nie jest zalecana. Akceptacja takiej metody jest ograniczona tylko do przypadków, w których pacjent nie może wnieść swojego wkładu lub gdy użycie pełnomocnictwa może być uzasadnione charakterem wydawanych orzeczeń.

Zaleca się przejrzyste raportowanie w wyznaczonym czasie wyników wszystkich pomiarów HRQoL. Jeżeli wyniki nie są (jeszcze) opublikowane, konieczne jest udostępnienie ich organom HTA, aby umożliwić krytyczną ocenę. Gdy zmiany w przeżyciu i HRQoL są łączone w jedną miarę wyników, taką jak QALY, należy wymagać osobnego zgłaszania zmian w przeżyciu i HRQoL oraz opisu metod łączenia pomiarów, aby umożliwić odrębne rozważenie obu punktów końcowych (36).

#### **4 Podsumowanie i wnioski**

W wyniku wyszukiwania odnaleziono wytyczne agencji i instytucji związanych z HTA z 16 państw. W większości wytycznych omawiano punkty końcowe w kontekście oceny analizy ekonomicznej.

Wskazuje się, że wybór konkretnego punktu końcowego zależy od takich czynników jak: populacja pacjentów spełniających kryteria kwalifikacji, stan zdrowia i cel leczenia. Punkty końcowe powinny odzwierciedlać przydatność interwencji w zakresie opieki zdrowotnej.

Spośród punktów końcowych definiowanych jako istotne klinicznie najczęściej wymieniano: punkty końcowe związane z przeżyciem, przebiegiem/nasileniem choroby, jakością życia związaną ze zdrowiem oraz występowaniem skutków ubocznych.

Zastosowanie zastępczych punktów końcowych (surogatów) jest zasadne, gdy: nie jest możliwe zastosowanie klinicznie istotnych punktów końcowych oraz wykazano wpływ tego punktu na klinicznie istotne punkty końcowe.

Wykorzystanie złożonych punktów końcowych jest zasadne, gdy uwzględnione składowe złożonego punktu końcowego są istotnymi punktami końcowymi powiązаныmi klinicznie. Uzasadnieniem łączenia różnych miar wyników jest zwiększenie mocy statystycznej badań, dla punktów końcowych, których częstotliwość występowania jest niska.

Odnalezione wytyczne agencji i instytucji związanych z HTA w zakresie definicji istotnych punktów końcowych, zastosowania zastępczych punktów końcowych oraz wykorzystania złożonych punktów końcowych są spójne z wytycznymi polskiej agencji HTA (AOTMiT) oraz EUnetHTA.

## 5 Bibliografia

1. Ustawa z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (Dz. U. z 2019 r., poz.1373.).
2. Ustawa z dnia 12 maja 2011 r. o refundacji leków, środków spożywczych specjalnego przeznaczenia żywieniowego oraz wyrobów medycznych (Dz. U. z 2019 r., poz. 784 z późn. zm.).
3. Wytyczne przeprowadzania Oceny Technologii Medycznych (HTA). Załącznik do Zarządzenia nr 40/2016 Prezesa Agencji Oceny Technologii Medycznych i Taryfikacji z dnia 13 września 2016 r. w sprawie wytycznych oceny świadczeń opieki zdrowotnej.
4. Polski Instytut Evidence Based Medicine. Słownik EBM. Słownik podstawowych pojęć. [http://ebm.org.pl/show.php?aid=15739&\\_tc=AD0C94700B6CF27881D8B67EAE8A883D](http://ebm.org.pl/show.php?aid=15739&_tc=AD0C94700B6CF27881D8B67EAE8A883D) [dostęp: 23.08.2019 r.]
5. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. NICE, Wielka Brytania 2014 (aktualizacja 2018).
6. Guyatt GH., et. al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. Journal of Clinical Epidemiology, 64 (2011): 395-400.
7. National Institute for Health and Care Excellence. Medicines optimisation: the safe and effective use of medicines to enable the best possible outcomes. NICE, Wielka Brytania 2015.
8. National Institute for Health and Care Excellence. Single technology appraisal: User guide for company evidence submission template. NICE, Wielka Brytania 2015.
9. National Institute for Health and Care Excellence. Evidence summaries: new medicines – Integrated process statement. NICE, Wielka Brytania 2013.
10. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. NICE, Wielka Brytania 2013.
11. Pharmaceutical Benefits Advisory Committee, Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee, Version 5.0, September 2016.
12. Belgian Health Care Knowledge Centre, Belgian Guidelines for Economic Evaluation and Budget Impact Analysis: second edition, KCE REPORT 183C, 2012.
13. Agency for Quality and Accreditation in Health Care, Department for Development, Research and Health Technology Assessment. The Croatian Guideline for Health Technology Assessment Process and Reporting, 1st ed Zagreb, February 2011.
14. State Institute for Drug Control. Cost-Effectiveness Analysis Critical Appraisal Procedure. Version: 3, Czechy 17.5.2017.
15. Lääkkeiden hintalautakunta, Läkemedelsprisnämnden (HILA). Preparing a health economic evaluation to be attached to the application for reimbursement status and wholesale price for a medicinal product. Application instructions TTS 31.1.2017.
16. Haute Autorité de Santé, HAS. A methodological guide. Choices in Methods for Economic Evaluation. Department of Economics and Public Health Assessment, HAS, Francja 2012.
17. Zorginstituut Nederland. Guideline for economic evaluations in healthcare, ZIN, Holandia 2016.
18. Ministry of Health, Welfare and Sport and College voor Zorgverzekeringen. Report. Dutch Assessment Procedures for the Reimbursement of Outpatient Medicines, MHWS, CVZ 2010.
19. Health Information and Quality Authority, Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland, December 2018.
20. Health Information and Quality Authority, Guidelines for the Economic Evaluation of Health Technologies in Ireland, 2018.
21. Pan-Canadian Oncology Drug Review, pCODR Procedures, June 2019.
22. Canadian Agency for Drugs and Technologies in Health, CADTH Methods and Guidelines, Guidelines for the Economic Evaluation of Health Technologies: Canada, 4th Edition, version 1.0, March 2017.



23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, General Methods Version 5.0, 10 July 2017.
24. Norwegian Institute of Public Health. Research gaps identified in systematic reviews and health technology assessments published by the Norwegian Knowledge Centre for the Health Services in 2015. Norwegian Institute of Public Health, Oslo 2017.
25. Pharmaceutical Management Agency, Guidelines for Funding Applications to PHARMAC, Amended in 2017
26. Pharmaceutical Management Agency, Prescription for Pharmacoeconomic Analysis, Methods for cost-utility analysis, version 2.2-2015
27. National Health Service Scotland, Scottish Medicine Consortium. Guidance to submitting companies for completion of New Product Assessment Form (NPAF). NHS Scotland, SMC 2019
28. NHS Scotland, Scottish Medicine Consortium. New Product Assessment Form for Ultra-Orphan Medicines. NHS Scotland, SMC 2019.
29. NHS Scotland, Scottish Medicine Consortium. Guidance to Submitting Companies for Completion of New Product Assessment Form (NPAF). Supplement for medicines for extremely rare conditions (ultra-orphan medicines). NHS Scotland, SMC 2019.
30. Swedish Agency on Health Technology Assessment and Assessment of Social Services. Assessment of methods in health care. A handbook. SBU 2018.
31. National Institute of Pharmacy and Nutrition. Professional Healthcare Guideline on the Methodology of Health Technology Assessment. English Version. The journal of the Hungarian pharmaceutical authority for doctors and pharmacists. Special issue. 2017.
32. European Network for Health Technology Assessment, EUnetHTA. Guideline. Endpoints used for Relative Effectiveness Assessment: Clinical Endpoints. Adapted version (2015).
33. European Network for Health Technology Assessment, EUnetHTA. Guideline. Endpoints used in Relative Effectiveness Assessment: Safety. Adapted version (2015).
34. European Network for Health Technology Assessment, EUnetHTA. Guideline. Endpoints used in Relative Effectiveness Assessment: Surrogate Endpoints. Adapted version (2015).
35. European Network for Health Technology Assessment, EUnetHTA. Guideline. Endpoints used for Relative Effectiveness Assessment: Composite Endpoints. Adapted version (2015).
36. European Network for Health Technology Assessment, EUnetHTA. Guideline. Endpoints used for Relative Effectiveness Assessment: Health-Related Quality of Life and Utility Measures. Adapted version (2015).